

# Hệ thống nhận diện biển số tự động

Võ Hữu Tuấn, Hồ Minh Đăng

Giảng viên hướng dẫn: Thầy Võ Hoài Việt, Thầy Phạm Minh Hoàng

Bộ môn Thị giác Máy tính, Khoa Công nghệ Thông tin

Trường Đại học Khoa học Tự nhiên - ĐHQGHCM

**Tóm tắt nội dung**—Automatic License Plate Recognition (ALPR) là công nghệ nhận diện biển số xe tự động trong các hệ thống giao thông thông minh (ITS), hỗ trợ giám sát giao thông và quản lý phương tiện. Bài báo cáo này trình bày một hệ thống ALPR có cấu trúc mô-đun (module) được xây dựng dựa trên nền tảng mô hình phát hiện đối tượng YOLOv5. Hệ thống bao gồm hai giai đoạn chính: phát hiện biển số và nhận diện ký tự. Một mô hình YOLOv5 được huấn luyện để phát hiện biển số xe và được chia làm 2 định dạng: biển số một dòng (1 line) và biển số hai dòng (2 line). Tiếp theo, một mô hình dựa trên mô hình YOLOv5 (base-on YOLO) được huấn luyện riêng để nhận diện các ký tự trên biển số được phát hiện. Hệ thống được đánh giá trên tập dữ liệu biển số từ nhiều khu vực (đa số là biển số 2 dòng ở Việt Nam). Kết quả sơ bộ cho thấy hệ thống có khả năng nhận diện chính xác các biển số 2 dòng của Việt Nam, kể cả trong điều kiện thiếu sáng. Tuy nhiên, hệ thống gặp một số sai sót nhỏ trong việc nhận diện các biển số 1 dòng ở nước ngoài (điển hình như biển số ở một số bang nước Mỹ). Nhìn chung, hệ thống ALPR này có hướng tiếp cận khá tốt trong việc nhận diện biển số và đã có kế hoạch cải tiến trong tương lai sắp tới.

## I. TỔNG QUAN

### A. Động lực

Trong bối cảnh càng ngày lượng phương tiện di chuyển ngày càng gia tăng nhanh chóng đã tạo ra áp lực rất lớn làm việc giám sát và quản lý giao thông bằng các phương pháp thủ công trở nên quá tải. Việc giám sát và quản lý phương tiện một cách tự động và chính xác theo thời gian thực trở thành nhu cầu cấp thiết. Trong đó phải kể đến mô hình nhận diện biển số xe tự động (ALPR) đóng vai trò ngày càng quan trọng trong các giải pháp giám sát và quản lý giao thông thông minh như:

- Giám sát giao thông (phạt nguội) trên các tuyến đường chính hay trên các tuyến đường cao tốc [1].
- Hệ thống thu phí bãi đỗ xe tự động và bán tự động ở các trung tâm thương mại hay các chung cư hay khu dân cư lớn [2].
- Hệ thống trạm thu phí cầu đường (BOT) và hệ thống thu phí không dừng (ETC) [3].

Xuất phát từ nhu cầu thiết yếu trong việc quản lý và giám sát giao thông thông minh, đặc biệt là vấn đề tự động nhận diện biển số xe trong bối cảnh hiện nay, chúng tôi quyết định nghiên cứu về các mô hình ALPR. Động lực chính khi lựa chọn hướng nghiên cứu này là mong muốn hiểu rõ cấu trúc và cách thức hoạt động của các mô hình này, từ đó xây dựng

một hệ thống ALPR riêng có khả năng nhận diện biển số xe một cách tự động (chủ yếu là nhận diện biển số xe Việt Nam).

### B. Phát biểu bài toán

Bài toán đặt ra là xây dựng một hệ thống có khả năng phát hiện biển số từ các ảnh đầu vào. Hệ thống cần xác định được vị trí biển số, phân đoạn và nhận diện chính xác các ký tự trên biển số đó. Các ảnh đầu vào được chụp trong điều kiện môi trường thực tế và theo nhiều góc chụp khác nhau.

### C. Hướng tiếp cận

Hướng tiếp cận chính được sử dụng trong báo cáo này là áp dụng các phương pháp hiện đại trong lĩnh vực thị giác máy tính, đặc biệt là sử dụng mô hình YOLOv5 giúp phát hiện đối tượng dựa trên deep learning với độ chính xác cao và hiệu suất thời gian thực.

## II. CÁC NGHIÊN CỨU LIÊN QUAN

Nhiều nghiên cứu đã được thực hiện trong lĩnh vực ALPR, tập trung vào các khía cạnh khác nhau của hệ thống.

**Các phương pháp trích xuất biển số**[4]: Các phương pháp này bao gồm phân tích biên cạnh, mô hình hóa xác suất, không gian con, phân tích blob, không gian màu và các thuộc tính hình học .... Các phương pháp dựa trên học sâu, đặc biệt là sử dụng **YOLO (You Only Look Once)**, đã cho thấy hiệu quả cao trong việc phát hiện biển số ....

**Các phương pháp phân đoạn ký tự**[5]: Các phương pháp phân đoạn ký tự được phân loại dựa trên các đặc trưng được sử dụng. Một phương pháp sử dụng lập trình động (DP) để phân đoạn các ký tự số trên biển số đã đạt tỷ lệ thành công 97,14% [5]. Ngoài ra các mạng phân đoạn được sử dụng như CRNET [6], DeepLabv2 hay ResNet-101.

**Các phương pháp nhận dạng ký tự**[5]: Các phương pháp nhận dạng ký tự có thể dựa trên giá trị pixel (ví dụ: so khớp mẫu) hoặc các đặc trưng được trích xuất. So khớp mẫu đơn giản và trực quan, nhưng dễ bị ảnh hưởng bởi sự thay đổi phông chữ, độ nghiêng, nhiễu và độ dày. Ngoài ra có thể sử dụng mạng học sâu CRNET[7] để xử lý và nhận diện ký tự. Một phương pháp khác là Character Timeseries Matching [8], để xử lý ký tự theo từng khung hình liên tiếp, giúp cải thiện độ chính xác khi nhận dạng trên video hoặc hình ảnh có biến dạng.

Một số công trình tập trung vào việc **xử lý các biển số đa dạng** từ nhiều quốc gia khác nhau .... Các yếu tố quan trọng để xử lý vấn đề này bao gồm góc xoay của biển số, số lượng dòng ký tự, loại ký tự chữ và số được sử dụng và định dạng ký tự. Các nghiên cứu gần đây cũng đã **khám phá sự tương tác giữa phân đoạn và nhận dạng ký tự**, sử dụng kiến thức tiên nghiệm về các ký tự để cải thiện phân đoạn và đưa kết quả nhận dạng trở lại quá trình phân đoạn.

### III. ĐỀ XUẤT

Dựa trên các bài nghiên cứu đã tìm hiểu, chúng tôi xin đề xuất mô hình hệ thống áp dụng phương pháp tiếp cận hiện đại là sử dụng và cải tiến mô hình YOLOv5 để giải quyết bài toán ALPR. Trình tự nhận diện biển số xe của hệ thống được trình bày tại Ảnh 1 Hệ thống mô hình sẽ được chia thành 2 mô-đun thực hiện 2 nhiệm vụ chuyên biệt như sau:

#### A. Mô-đun phát hiện và phân loại biển số

Tận dụng điểm mạnh của mô hình YOLOv5 về việc phát hiện và nhận diện các vật thể có kích thước vừa và nhỏ trong ảnh, ở mô-đun này huấn luyện YOLOv5 truyền thống[9] để xác định vị trí của biển số, tách hình ảnh biển số và phân loại chúng theo 2 loại là: biển số 1 dòng (1\_line) và biển số 2 dòng (2\_line) - phổ biến ở Việt Nam.

#### B. Mô-đun phân đoạn và nhận dạng ký tự

Sau khi phát hiện biển số xe, bước kế tiếp là nhận diện các ký tự trên biển số. Ở đây chúng tôi tinh chỉnh cấu trúc mô hình YOLOv5 một chút để phù hợp và thích nghi tốt hơn với 2 loại biển số ở tập dataset. Tổng quan cấu trúc của mô-đun này được trình bày tại bảng 1.

Bảng 1: Tổng quan cấu trúc Mô-đun nhận dạng ký tự [8]

Module index	Connected from	Module	Input size	Output size
0	-	Conv	$3 \times 128 \times 128$	$16 \times 64 \times 64$
1	0	Conv	$16 \times 64 \times 64$	$32 \times 32 \times 32$
2	1	C3	$32 \times 32 \times 32$	$32 \times 32 \times 32$
3	2	Conv	$32 \times 32 \times 32$	$64 \times 16 \times 16$
4	3	C3	$64 \times 16 \times 16$	$64 \times 16 \times 16$
5	4	Conv	$64 \times 16 \times 16$	$128 \times 8 \times 8$
6	5	C3	$128 \times 8 \times 8$	$128 \times 8 \times 8$
7	6	Conv	$128 \times 8 \times 8$	$256 \times 4 \times 4$
8	7	SPP	$256 \times 4 \times 4$	$256 \times 4 \times 4$
9	8	C3TR	$256 \times 4 \times 4$	$256 \times 4 \times 4$
10	9	Conv	$256 \times 4 \times 4$	$128 \times 4 \times 4$
11	10	Upsample	$128 \times 4 \times 4$	$128 \times 8 \times 8$
12	[6,11]	Concatenate	$[128 \times 8 \times 8]$ $[128 \times 8 \times 8]$	$256 \times 8 \times 8$
13	12	C3	$256 \times 8 \times 8$	$128 \times 8 \times 8$
14	13	Conv	$128 \times 8 \times 8$	$64 \times 8 \times 8$
15	14	Upsample	$64 \times 8 \times 8$	$64 \times 16 \times 16$
16	[4,15]	Concatenate	$[64 \times 16 \times 16]$ $[128 \times 8 \times 8]$	$128 \times 16 \times 16$
17	16	C3	$128 \times 16 \times 16$	$64 \times 16 \times 16$
18	17	Detect	$64 \times 16 \times 16$	-

Cụ thể, mô hình phát hiện ký tự được xây dựng dựa trên YOLOv5 với việc tích hợp khối tự chú ý đa đầu (multi-head

self-attention) nhằm tăng khả năng trích xuất đặc trưng của các ký tự trên biển số. Kích thước đầu vào được đặt ở mức  $128 \times 128$ , đảm bảo sự cân bằng giữa tốc độ xử lý và độ chính xác nhận diện. Do các ký tự có kích thước tương đối lớn so với tổng thể biển số, chúng tôi loại bỏ các đầu phát hiện (detection head) cho các đặc trưng nhỏ và trung bình, đồng thời giảm một nửa số lượng bộ lọc ở mỗi lớp. Điều này giúp mô hình hoạt động nhanh hơn mà không ảnh hưởng đáng kể đến độ chính xác nhận diện. Việc sử dụng C3TR giúp mô hình có khả năng tập trung tốt hơn vào các đặc trưng quan trọng của từng ký tự, cải thiện độ chính xác nhận diện trong môi trường thực tế với các điều kiện ánh sáng và góc chụp khác nhau.

1) *Trích xuất ký tự của biển 1 dòng*: Sau khi phát hiện và nhận diện các ký tự trên biển số, đối với biển số thuộc loại 1 dòng, mô-đun sẽ sắp xếp theo thứ tự tăng dần của trục x và gộp các ký tự đã nhận diện lại và trả về kết quả tổng hợp duy nhất cho biển số.

2) *Trích xuất ký tự của biển 2 dòng*: Đối với biển số thuộc loại 2 dòng, sau khi phát hiện và nhận diện các ký tự, mô-đun sẽ áp dụng phương pháp hồi quy tuyến tính để tính toán và tìm ra một đường thẳng phân cắt giữa hai dòng ký tự. Cụ thể, ta tính toán tọa độ trung tâm của các bounding box (bbox) của mỗi ký tự, được xác định bởi các giá trị  $center_x$  và  $center_y$ :

$$center_x = \frac{x_{\min} + x_{\max}}{2}, \quad center_y = \frac{y_{\min} + y_{\max}}{2} \quad (1)$$

Sau đó, mô-đun sẽ áp dụng phương pháp hồi quy tuyến tính để xác định đường phân chia giữa các ký tự ở hai dòng. Phương trình của đường thẳng phân cắt có dạng:

$$y = ax + b \quad (2)$$

Trong đó: -  $a$  là hệ số góc của đường thẳng (độ dốc), -  $b$  là giao điểm của đường thẳng với trục  $y$ .

Các giá trị  $a$  và  $b$  được tính thông qua công thức bình phương tối thiểu:

$$a = \frac{n \sum_{i=1}^n (x_i y_i) - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \quad (3)$$

$$b = \frac{\sum_{i=1}^n y_i - a \sum_{i=1}^n x_i}{n} \quad (4)$$

Trong đó:

- $x_i$  và  $y_i$  là các tọa độ của điểm  $i$  (tức là các giá trị  $center_x$  và  $center_y$  của các ký tự).
- $n$  là số lượng ký tự đã nhận diện.

Sau khi tính toán được giá trị của  $a$  và  $b$ , mô-đun sẽ xác định được đường thẳng phân chia các ký tự vào hai nhóm: nhóm ký tự ở dòng trên và nhóm ký tự ở dòng dưới. Kết quả là mô-đun sẽ sắp xếp theo thứ tự tăng dần và trả về các ký tự phân tách chính xác cho từng dòng và sau đó kết hợp chúng lại để cho ra kết quả cuối cùng.

Bằng cách kết hợp 2 mô-đun này, hệ thống đề xuất sẽ có khả năng xử lý nhiều loại biển số, đặc biệt là các biển số xe 2 dòng của Việt Nam.



Ảnh 1: Trình tự hoạt động của mô hình hệ thống

#### IV. THỰC NGHIỆM

##### A. Tập dữ liệu

Hệ thống được thử nghiệm và đánh giá trên tập dữ liệu được tổng hợp các hình ảnh biển số từ tập dữ liệu **UFPR-ALPR** [10] và tập dữ liệu **Vietnamese License Plate** [11]. Đảm bảo ảnh đầu vào bao gồm cả 2 loại biển số và được dưới nhiều góc độ và điều kiện môi trường khác nhau.

##### B. Độ đo (metric)

Tập trung vào **độ chính xác (accuracy)**, độ đo cơ bản và phổ biến trong việc đánh giá các mô hình thuộc lĩnh vực thị giác máy tính. Độ chính xác cho biết tỷ lệ dự đoán đúng trên tổng số mẫu trong tập kiểm tra. Công thức tính độ chính xác được tính toán như sau:

$$\text{Accuracy} = \frac{\text{Số lượng dự đoán đúng}}{\text{Tổng số mẫu}} \quad (5)$$

Độ chính xác là một chỉ số dễ hiểu và trực quan. Nó phản ánh khả năng tổng quát của mô hình trong việc phân loại và nhận diện biển số xe của mô hình hệ thống này.

##### C. Thử nghiệm

1) **Mục tiêu:** Xây dựng một hệ thống nhận diện biển số qua hình ảnh với khả năng nhận diện chính xác cho các biển số được chụp trong điều kiện tốt và cho phép sai sót một cách tương đối với các hình ảnh biển số khó nhận diện do chất lượng ảnh quá thấp (bị nhòe, nhiễu nặng) hoặc bị che khuất quá nhiều dẫn đến sai lệch trong dự đoán. Và từ kết quả thí

nghiệm mang lại có thể xác định được ưu nhược điểm của mô hình để đề ra giải pháp cải tiến khả năng nhận diện tốt hơn với mong muốn hệ thống có thể phát hiện biển số trong các điều kiện môi trường xấu.

2) **Setup:** Thí nghiệm được thực hiện trên phiên bản Python 3.10, sử dụng Pytorch framework chạy trên hệ điều hành Ubuntu 24.04.2 với CPU Intel core i5 và 16GB RAM. Cấu trúc của hệ thống bao gồm 2 mô-đun như đã được trình bày ở Mục III.

3) **Kết quả:** Mô hình nhận diện đúng **1432** biển số trên **1800** mẫu kiểm tra, **độ chính xác** đạt được là **79.50%**. Điều này cho thấy mô hình có khả năng nhận diện khá tốt trong môi trường thực tế với độ chính xác khá cao (gần 80%).

4) **Nhận xét:** Dựa trên kết quả tổng quát đạt được và qua một số mẫu kiểm tra được rút trích ở Ảnh 2 và Ảnh 3 cho thấy mô hình hoạt động ổn định trên nhiều loại biển số khác nhau về màu sắc (biển nền trắng, xanh hay với cả biển đỏ và biển vàng), biển số 1 dòng và 2 dòng.

Mô hình nhận diện chính xác trong điều kiện ảnh được chụp với ánh sáng tốt, rõ nét. Một số ảnh mờ, nhiễu nhẹ vẫn được mô hình nhận diện tốt do vẫn giữ độ tương phản cao giữa màu nền biển và màu của ký tự.

Ở các trường hợp sai sót 3, điểm chung để nhận thấy nhất đó là ảnh bị nhiễu và mờ nặng làm các ký tự khó phân biệt, hoặc do góc chụp nghiêng hoặc biển số nhỏ khiến mô hình bị mất mát thông tin của ký tự.

Đặc biệt vì mô hình chưa được huấn luyện nhận dạng ký tự "Đ" trong bảng chữ cái Tiếng Việt nên dẫn đến sai sót trong



Ảnh 2: Một số kết quả nhận diện đúng

các mẫu chứa ký tự này (ví dụ biển số 79MD104296, mô hình nhận dạng ký tự "Đ" thành "D"). Đây là một thiếu sót khá lớn trong việc huấn luyện mô hình và cần cải thiện đầu tiên.

Với độ chính xác gần 80% là một mức khá tốt nhưng chưa đủ mức "tin cậy" để triển khai thực tế (thường kỳ vọng ít nhất từ 90% trở lên cho hệ thống giám sát). Nhìn chung, mô hình hoạt động khá tốt và có tiềm năng cải thiện để xử lý tốt hơn các trường hợp khó.

## V. HƯỚNG CẢI TIẾN TRONG TƯƠNG LAI

### A. Xây dựng tập dữ liệu tốt hơn

Vì tập dữ liệu có quy mô còn khá nhỏ (chỉ có 1800 mẫu) nên cần mở rộng lớn hơn để kết quả thu được khách quan hơn. Đặc biệt là việc bổ sung các mẫu dữ liệu về ký tự "**Đ**" vào tập huấn luyện là rất cần thiết. Vì trong thực nghiệm mô hình đã nhận diện sai các mẫu chứa ký tự này

### B. Áp dụng các phương pháp tiền xử lý ảnh

Đối với các trường hợp khó như ảnh bị mờ nhiều nặng hay góc chụp nghiêng cần áp dụng các phương pháp tiền xử lý để tăng chất lượng ảnh đầu vào giúp tăng khả năng nhận diện của mô hình.

Các phương pháp tiền xử lý được cân nhắc:

- Cân bằng độ tương phản (histogram equalization)
- Khử nhiễu và làm nét ảnh (denoising và sharpening)
- Hiệu chỉnh góc (rotation correction)

Việc khử nhiễu, làm nét ảnh hay cân bằng độ tương phản giúp tăng chất lượng hình ảnh và việc áp dụng phương pháp hiệu chỉnh góc đưa biển số về thẳng hàng giúp dễ tách dòng cũng như nhận diện ký tự dễ hơn

### C. Áp dụng phương pháp hậu xử lý

Cần nhắc sử dụng phương pháp "**knowledge domain**": áp dụng thêm các quy tắc của biển số Việt Nam. Ví dụ như áp dụng các tiền tố đặc biệt của biển số xe Việt Nam như



Ảnh 3: Một số kết quả nhận diện sai

"80A" cho các biển xe công vụ (biển đỏ) để xác thực kết quả nhận diện biển số.

## TÀI LIỆU

- [1] Ángel Díaz, Rachel Levinson-Waldman, "Automatic License Plate Readers: Legal Status and Policy Recommendations for Law Enforcement Use", 10 September 2020. Brennan Center for Justice
- [2] M. Y. I. Idris et al., "Automatic Parking Management System and Parking Fee Collection Based on Number Plate Recognition", January 2012 . Researchgate: 281060377
- [3] Y. Sun et al., "Electronic Toll Collection (ETC) Deployment on Highways", 4 Jun 2024. papers.ssrn.com
- [4] R. Al-Batat, A. Angelopoulou, S. Premkumar, J. Hemanth, and E. Kapetanios, "An End-to-End Automated License Plate Recognition System Using YOLO Based Vehicle and License Plate Detection with Vehicle Classification,"extitSensors, vol. 22, no. 9477, 2022. DOI: 10.3390/s22239477.
- [5] S. Du, M. Ibrahim, M. Shehata, and W. Badawy, "Automatic License Plate Recognition (ALPR): A State-of-the-Art Review,"extitIEEE Trans. Circuits Syst. Video Technol., vol. 23, no. 2, pp. 311-325, 2013. DOI: 10.1109/TCSVT.2012.2203741.
- [6] Rayson Laroca , Evair Severo , Luiz A. Zanlorensi , Luiz S. Oliveira, Gabriel Resende Gonc alves , William Robson Schwartz and David Menotti, "A Robust Real-Time Automatic License Plate Recognition Based on the YOLO Detector", 28 Apr 2018. arXiv: 1802.09567

- [7] Rayson Laroca, Luiz A. Zanlorensi, Gabriel R. Gonçalves, Eduardo Todt, William Robson Schwartz, David Menotti, "An efficient and layout-independent automatic license plate recognition system based on the YOLO detector". DOI: 10.1049/itr2.12030
- [8] Quang Huy Che, Tung Do Thanh, Cuong Truong Van, "Character Time-series Matching For Robust License Plate Recognition", 13 Sep 2023. ARXIV: 2307.11336
- [9] Glenn-jocher, UltralyticsAssistant, "YOLOv5 in PyTorch". [github.com](https://github.com)
- [10] R. Laroca, E. Severo, L. A. Zanlorensi, L. S. Oliveira, G. R. Gonçalves, W. R. Schwartz, D. Menotti, "A Robust Real-Time Automatic License Plate Recognition Based on the YOLO Detector," in 2018 International Joint Conference on Neural Networks (IJCNN), pp. 1-10, July 2018. [Webpage] [IEEE Xplore] [PDF] [BibTeX] [NVIDIA News]
- [11] Namu Desu, "Vietnamese License Plate OCR dataset", Nov 2024. [kaggle.com](https://www.kaggle.com)
- [12] Nguyen Mai (@gawrgura), "Tổng hợp kiến thức từ YOLOv1 đến YOLOv5", 24 June 2022. [viblo.asia]