NHẬN DIỆN VÀ PHÂN LOẠI   
BIỂN BÁO GIAO THÔNG

Nguyễn Anh Tuấn   
dept. Computer Vision   
University of Science, VNU   
Ho Chi Minh City, Viet Nam  
19127614@student.hcmus.edu.vn

Nguyễn Hoàng Vũ  
dept. Computer Vision   
University of Science, VNU   
Ho Chi Minh City, Viet Nam  
19127632@student.hcmus.edu.vn

*Abstract*— Phát hiện và nhận dạng biển báo giao thông nắm một vai trò quan trọng trong các hệ thống giao thông thông minh hiện nay, Đồng thời còn đóng góp vào các ứng dụng thực tế như hệ thống xe tự hành, hệ thống hỗ trợ giao thông,.. Báo cáo này sẽ xây dựng một hệ thống nhận diện biển báo giao thông bằng cách xây dựng một mạng CNN đơn giản dựa trên model mạng học sâu YOLOv4.

Keywords: Traffic signs, detection, recognition, Advance Driver Assistance System, CNN

# **GIỚI THIỆU**

Với sự phát triển không ngừng nghỉ của xã hội cũng như sự tiến bộ của khoa học, nhu cầu di chuyển của con người nay ngày càng được lưu tâm. Từ việc đi bộ, cưỡi động vật,... cho đến di chuyển bằng các xe cơ giới. Việc di chuyển bằng xe cơ giới này tạo nên một hệ thống giao thông vô cùng phức tạp.Do đó, “biển báo giao thông” là một trong những thứ đóng vai trò rất quan trọng trong việc truyền tải các nội dung quan trọng một cách nhanh nhất cho các người sử dụng phương tiện (ví dụ: biển báo cấm vượt, quẹo, biển báo giới hạn tốc độ, các biển báo thông báo nguy hiểm,...). Việc này đồng thời mạng lại sự an toàn, đồng thời đảm bào các hệ thống giao thông, cầu đường hoạt động một cách trơn tru khi mà các người lái đều có thông tin về đoạn đường mình di chuyển.

Tuy nhiên, thời đại 4.0 sinh ra một hệ thống cao cấp hơn của giao thông: các hệ thống giao thông thông minh (intelligent transportation system) bao gồm những hệ thống nhúng, giám sát, hỗ trợ giao thông và đồng thời trong tương lai cũng có thể suất hiện một hệ thống nữa đó là: hệ thống xe tự hành. Hai hệ thống này đề cao tính “tự động” trong việc quản lý, di chuyển phương tiện giao thông, do đó, việc khiến cho các hệ thống trên nhận biết được biển báo giao thông là một điều vô cùng cần thiết.

Nhờ vậy, bài toán nhận diện và phân loại biển báo giao thông được ra đời và đang là đề tài nóng hổi cho các nhà nghiên cứu thị giác máy tính và trí tuệ nhân tạo tìm tòi, nghiên cứu. Bài báo cáo này sẽ nêu lên phương pháp xây dựng một hệ thống nhận diện biển báo thông qua một mạng neuron tích chập CNN sử dụng model YOLOv4, một vài ví dụ thực tế cũng như một vài cải tiến nhỏ để phát triển hệ thống.

# **TỔNG QUAN**

## VẤN ĐỀ VÀ BÀI TOÁN (PROBLEM STATEMENT)

Bài toán phân loại biển báo giao thông cũng giống như các bài toán phân loại khác, ẩn trong nó là 2 bài toán nhỏ:

Thứ nhất: bài toán dò tìm đối tượng (Object Detection), ở bước này, ta sẽ “tìm” vật thể, ở đây là các biển báo giao thông. “Tìm” ở đây chính là xác định vị trí tọa độ trên ảnh đồng thời xác định kích thước của vật thể đó.

Thứ hai: bài toán nhận diện (Object Recognition), bài toán này hướng đến việc sẽ đi phân lớp các vật thể vào các lớp của nó. Ở đây, chúng ta sẽ phân loại xem các biển báo mà ta tìm được ở bài toán 1 là biển báo loại gì.

## CÁC NGHIÊN CỨU LIÊN QUAN (RELATED WORK)

Từ trước tới nay, đã có rất nhiều nghiên cứu về việc nhận diện được biển báo giao thông:

Việc phát hiện biển báo giao thông thường dựa trên các đặc trưng hình dạng và màu sắc.

Các biển báo giao thông đa phần đều tuân theo một bảng màu đã được quy định. Do đó, việc sử dụng màu như một đặc trưng để dò tìm các biển báo là một trong những hướng đi lâu đời của bài toán. Các phương pháp phát hiện biển báo giao thông dựa trên màu sắc (Color) là các phương pháp lâu đời và dễ hiểu **[2].** Tuy nhiên thì do hệ màu RGB khá nhạy cảm với các yếu tố khác như: ánh sáng, nhiễu, blur,... Đã có nhiều nghiên cứu đã được đề ra như: sử dụng hệ màu HSV**[6]**, YUV, HSI,… Dù đã cải tiến được đôi chút, song, các đặc trưng màu này vẫn quá bị ảnh hưởng bởi các yếu tố khác như: thời tiết, ánh sáng, …

Song song cùng với đặc trưng màu, đặc trưng dáng (Shape) cũng được dùng để nhận diện biển báo. Các biển báo thường có hình dáng cố định đặc trưng như: Tam giác, vuông, lục giác, tròn,… Dựa trên việc này, nhiều nghiên cứu đã loại bỏ hoàn toàn đặc trưng màu và thay vào đó là các đặc trưng dáng để nhận dạng **[7].**

Ở những thời kỳ đầu của bài toán, việc phân loại các đặc trưng này thường được dựa trên các thuật toán phân lớp như: Random Forest **[1]**, SVM **[2]**, KNN, …Các phương pháp trên đều có thể hoạt động, nhưng vẫn có nhiều hạn chế.

Bài toán này trở nên đột phá khi mà được kết hợp với mạng neuron tích chập CNN. Từ khi áp dụng CNN vào, đã có vô số các nghiên cứu để đưa ra các model, các thuật toán phù hợp với việc phân loại biển báo giao thông. **[4]** sử dụng model VGG đồng thời cải tiến nó thành một model mới có tên IVGG (improved VGG), tuy nhiên phương pháp này lại có nhược điểm là có tỉ lện nhận dạng thấp với các ảnh có nền tối, mờ. **[5]** Tự xây dựng một mô hình CNN đơn giản, đồng thời tinh chỉnh các mô hình như Faster RCNN và YOLOv4, báo cáo cũng cho thấy được sự so sánh giữa hiệu năng các phiên bản YOLO từ v1 – v4 với nhau. **[8]** xây dựng một quy trình dựa trên phân loại đặc trưng màu , sừ dụng HOG kết hợp với mạng CNN, để phát hiện và nhận dạng các biển báo giao thông, giúp đạt được độ chính xác phân loại và tốc độ tính toán tốt hơn. Tuy nhiên dù mạng lại kết quả rất khả quan trong độ chính xác khi nhận dạng biển báo, việc sử dụng các thuật toán deep learning khiến cho thời gian tính toán lâu hơn, dẫn đến một vấn đề khá lớn trong việc áp dụng vào real-time, thực tế. Vấn đề này vẫn luôn được nghiên cứu cũng như đưa ra các giải pháp nhằm cải thiện xuyên suốt thời gian gần đây.

# **PHƯƠNG PHÁP**

## Sơ đồ tổng quát

## Tập dữ liệu (Dataset Collection)

Sử dụng dataset là: German Traffic Sign Recognition Benchmark, [GTSRB](https://www.kaggle.com/meowmeowmeowmeowmeow/gtsrb-german-traffic-sign?select=Train).

Được sử dụng lần đầu tại IJCNN năm 2011,bộ dữ liệu này bao gồm khoảng 51839 hình ảnh được lưu trữ dưới định dạng png chia làm 43 lớp khác nhau theo chủ đề biển báo giao thông. Trong đó có 39,209 được dùng là ảnh dùng cho training set và 12,630 ảnh dùng để test và thẩm định (validation). Các ảnh có kích thước từ khoảng 15x15 cho đến 250x250 pixels .Các ảnh trong tập dữ liệu đều có điều kiện ánh sáng, khoảng cách, thời tiết phong phú, đa dạng, giúp cho việc tranning và testing tốt hơn. Tập dữ liệu có một thư mục train chứa các hình ảnh bên trong mỗi lớp và một thư mục test chứa các hình ảnh dùng để thử nghiệm mô hình

## Tiền xử lý và phân chia dữ liệu

* Đây là bước đầu tiên của hệ thống nhận diện: load dữ liệu và xử lý dữ liệu training. Chia tập dữ liệu ra thành các tập như: Training Set – dùng để huấn luyện máy học.
* Testing Set - là tập dữ liệu dùng để test sau khi máy đã học xong
* Validation Set - test cũng giống như tập training set, nó cũng bao gồm các cặp giá trị input và ouput tương ứng. Nhưng nó lại khác training set ở chỗ, nó được sử dụng để kiểm thử độ chính xác của mô hình máy học trong quá trình huấn luyện

Lưu ý: Testing Set được dùng để kiểm thử sau quá trình huấn luyện, còn Validation Set được sử dụng để kiểm thử trong quá trình huấn luyện.

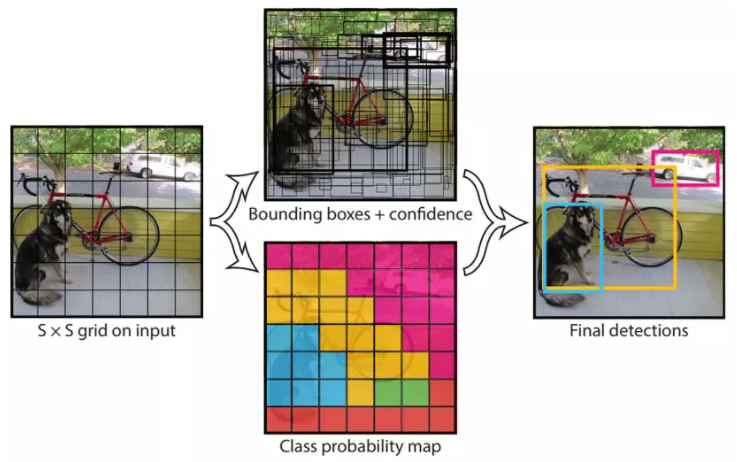
Tiếp đó sẽ là bước xử lý dữ liêu, cơ bản làm bước dùng để biến đổi ảnh đầu vào thành một ảnh phù hợp hơn, bằng các cách như resize ảnh, blur ảnh, ...

## Thuật toán YOLO

### Sơ lược về YOLO

Yolo là một mô hình mạng học sâu CNN, dùng để phát hiện, nhận dạng, phân loại đối tượng, từ việc sử dụng các lớp convolutional layers và lớp connected layers. Trong đó các convolutional layers sẽ trích xuất ra các feature của ảnh, còn full-connected layers sẽ dự đoán ra xác suất đó và tọa độ của đối tượng.

Kiến trúc YOLO coi bài toán phát hiện vật như một bài toán regression. Từ input là ảnh đầu vào, qua một mạng gồm các lớp convolution, pooling và fully connected là có thể ra được output. Kiến trúc này có thể được tối ưu để chạy trên GPU với một lần forward pass, và vì thế đạt được tốc độ rất cao.

Ý tưởng chính của YOLOv1 là chia ảnh thành một lưới các ô (grid cell) với kích thước SxS ( 3x3, 5x5, 7x7 ...) Với mỗi grid cell, mô hình sẽ đưa ra dự đoán cho bounding box

### Thuật toán YOLOv4

Intersection over Union là chỉ số đánh giá được sử dụng để đo độ chính xác của Object detector trên tập dữ liệu cụ thể. Chúng ta hay gặp chỉ số này trong các Object Detection Challenge, dạng nhưPASCAL VOC challenge.

Sẽ thường thấy IoU được sử dụng để đánh giá performance của các bộ object detector HOG + Linear SVM và Convolutional Neural Network (R-CNN, FasterR-CNN, YOLO, v.v...). Tuy nhiên đối với IoU thì sử dụng thuật toán nào để đưa ra các prediction không quan trọng.

IoU đơn giản chỉ là một chỉ số đánh giá. Mọi thuật toán có khả năng predict ra các bounding box làm output đều có thể được đánh giá thông qua IoU.

Để áp dụng được IoU để đánh giá một object detector bất kì ta cần:

Những ground-truth bounding box (bounding box đúng của đối tượng, ví dụ như bounding box của đối tượng được khoanh vùng và đánh nhãn bằng tay sử dụng trong tập test.)

Những predicted bounding box được model sinh ra.

Miễn là có hai tập bên trên, ta đều có thể sử dụng được IoU.

Nhìn vào công thức này, bạn sẽ thấy IoU đơn giản là một tỉ lệ. Ở tử số ta tính toán area of overlap - diện tính phần chồng lên nhau giữa predicted bounding box và ground-truth bounding box. Phần mẫu số là area of union - diện tích phần hợp - hay đơn giản hơn là diện tích mà hai bounding box này đang chiếm. Chia diện tích phần chồng (giao) cho diện tích phần hợp sẽ thu được giá trị mà ta mong muốn - Intersection over Union (IoU).

Trước khi đi sâu hơn về IoU, có thể có người trong chúng ta sẽ tự hỏi những mẫu ground-truth được lấy ở đâu?. Như đã nói ở trên, ground-truth bounding box được "khoanh vùng và đánh nhãn bằng tay".

## Kiểm thử mô hình

# **TÍNH TOÁN VÀ THỰC NGHIỆM**

Mẫu: German Traffic Sign.

# **KẾT LUẬN**

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

## Authors and Affiliations

**The template is designed for, but not limited to, six authors.** A minimum of one author is required for all conference articles. Author names should be listed starting from left to right and then moving down to the next line. This is the author sequence that will be used in future citations and by indexing services. Names should not be listed in columns nor group by affiliation. Please keep your affiliations as succinct as possible (for example, do not differentiate among departments of the same organization).

### For papers with more than six authors: Add author names horizontally, moving to a third row if needed for more than 8 authors.

### For papers with less than six authors: To change the default, adjust the template as follows.

#### Selection: Highlight all author and affiliation lines.

#### Change number of columns: Select the Columns icon from the MS Word Standard toolbar and then select the correct number of columns from the selection palette.

#### Deletion: Delete the author and affiliation lines for the extra authors.

## Identify the Headings

Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Component heads identify the different components of your paper and are not topically subordinate to each other. Examples include Acknowledgments and References and, for these, the correct style to use is “Heading 5”. Use “figure caption” for your Figure captions, and “table head” for your table title. Run-in heads, such as “Abstract”, will require you to apply a style (in this case, italic) in addition to the style provided by the drop down menu to differentiate the head from the text.

Text heads organize the topics on a relational, hierarchical basis. For example, the paper title is the primary text head because all subsequent material relates and elaborates on this one topic. If there are two or more sub-topics, the next level head (uppercase Roman numerals) should be used and, conversely, if there are not at least two sub-topics, then no subheads should be introduced. Styles named “Heading 1”, “Heading 2”, “Heading 3”, and “Heading 4” are prescribed.

## Figures and Tables

#### Positioning Figures and Tables: Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation “Fig. 1”, even at the beginning of a sentence.

1. Table Type Styles

| Table Head | Table Column Head | | |
| --- | --- | --- | --- |
| Table column subhead | Subhead | Subhead |
| copy | More table copya |  |  |

1. Sample of a Table footnote. (*Table footnote*)
2. Example of a figure caption. (*figure caption*)

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity “Magnetization”, or “Magnetization, M”, not just “M”. If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write “Magnetization (A/m)” or “Magnetization {A[m(1)]}”, not just “A/m”. Do not label axes with a ratio of quantities and units. For example, write “Temperature (K)”, not “Temperature/K”.

##### Acknowledgment *(Heading 5)*

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

##### **References**

1. Jack Greenhalgh and M. Mirmehdi, [“Traffic sign recognition using MSER and Random Forests”](https://ieeexplore.ieee.org/abstract/document/6333901)
2. David Soendoro and Iping SuprianaTraffic, [Sign Recognition with Color-based Method, Shape-arc Estimation and SVM](https://ieeexplore.ieee.org/document/6021584)
3. Faming Shao, Xinqing Wang, Fanjie Meng, Ting Rui, Dong Wang, and Jian Tang, [Real-Time Traffic Sign Detection and Recognition Method Based on Simplified Gabor Wavelets and CNNs](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6210476/)
4. Shuren Zhou, Wenlong Liang, Junguo Li and Jeong-Uk Kim, [Improved VGG Model for Road Traffic Sign Recognition](https://www.techscience.com/cmc/v57n1/22953)
5. Njayou Youssouf, [Traffic Sign Detection and Recognition with](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4012734)

[FasterRCNN and YOLOV4](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4012734)

1. Hasan Fleyeh, Syed Omer Gilani[, Road sign detection and recognition using fuzzy artmap: A case study swedish speed-limit signs.](https://www.researchgate.net/publication/220909199_Road_sign_detection_and_recognition_using_fuzzy_artmap_A_case_study_swedish_speed-limit_signs)
2. Pavel Paclíck, Jana Novovicova, [Road Sign Classification without Color Information](https://www.semanticscholar.org/paper/Road-Sign-Classification-without-Color-Information-Pacl%C3%ADk-Novovicov%C3%A1/a0d1f3078208fb101e66d54765f86aeb8d606678)
3. Ali Youssef, Dario Albani, Daniele Nardi, and Domenico D. Bloisi, [Fast Traffic Sign Recognition Using Color Segmentation and Deep Convolutional Networks](https://www.semanticscholar.org/paper/Fast-Traffic-Sign-Recognition-Using-Color-and-Deep-Youssef-Albani/d1cc94c082313dc31f8885f003334cb4283c3d04)