

# Proposal for Automatic License Plate Recognition (ALPR)

1<sup>st</sup> Phùng Tiến Hào - 1612174

*Đại học Khoa học tự nhiên*

*Bộ môn Thị Giác Máy Tính*

Thành phố Hồ Chí Minh, Việt Nam

Email: tienhaophung@gmail.com

2<sup>nd</sup> Võ Quốc Huy - 1612269

*Đại học Khoa học tự nhiên*

*Bộ môn Thị giác máy tính*

Thành phố Hồ Chí Minh, Việt Nam

Email: voquochuy304@gmail.com

3<sup>rd</sup> Trần Nhật Huy - 1612272

*Đại học Khoa học tự nhiên*

*Bộ môn Thị giác máy tính*

Thành phố Hồ Chí Minh, Việt Nam

Email: nhathuy13598@gmail.com

**Abstract**—Bài toán Automatic license plate recognition (ALPR), hay còn có một số tên gọi khác như LPR, ANPR, VLPR có ý nghĩa quan trọng trong các ứng dụng liên quan đến ảnh đường xá. Công dụng của nó là lấy ra được thông tin từ các biển số xe từ một ảnh hay chuỗi các ảnh. Chất lượng ảnh thu vào là yếu tố chính ảnh hưởng đến sự thành công của ALPR. ALPR được xem là thành công khi nó trích xuất nhanh chóng và chính xác các thông tin biển số xe dưới những điều kiện môi trường khác nhau như trong nhà, ngoài trời, buổi sáng/tối. Ngoài ra, một số yếu tố khác cũng ảnh hưởng đến độ chính xác của ALPR là bóng, độ chiếu sáng, sự khác nhau về dáng của ảnh biển số,... Và sự đa dạng về biển số xe (font, ngôn ngữ, background,...) của từng quốc gia, địa phương cũng cần rất được quan tâm đến.

**Index Terms**—License plate detection, License plate recognition, Automatic license plate recognition (ALPR), Support vector machine (SVM), Optical character recognition (OCR).

## I. INTRODUCTION

License plate recognition (LPR) là một trong những research topic quan trọng về hệ thống vận chuyển thông minh, nó dần trở nên ngày càng được chú ý đến trong thời gian gần đây. Cùng với sự phát triển của Computer Vision, LPR đã được đưa ra ứng dụng thực tế như thu phí tự động, kiểm soát bãi đỗ xe, giám sát tình hình giao thông trên đường phố,... LPR là sự kết hợp của Object Detection, Image Processing, và Pattern Recognition. Sự đa dạng của loại biển số xe và môi trường là những thách thức trong việc phát hiện và nhận diện biển số xe.

LPR gồm 3 bước chính: License Plate Detection, License Plate Segmentation và Character Recognition. Bước thứ nhất giúp phát hiện vùng nào trên ảnh là biển số xe dựa trên đặc trưng viền, màu sắc hay sự tồn tại các ký tự,... Sau khi có được ảnh (patch) các biển số xe, sẽ tiến tới thực hiện phân đoạn và trích xuất được các vùng ký tự riêng biệt. Bước cuối cùng là nhận dạng được từng ký tự vừa được trích xuất bằng các phương pháp như neural network,... Để có được kết quả tốt nhất thì cả 3 bước trên đều phải được thực hiện kỹ càng.

Việc localization và segmentation biển số xe trong ảnh một cách chính xác mang tính quyết định đến toàn bộ hệ thống LPR. Có hai giải pháp chính: một là dựa trên đặc trưng màu và hai là dựa trên đặc trưng cạnh của biển số xe.

Các ký tự trong ảnh biển số xe sau đó mới được phân ra (segment) riêng rẽ bằng các phương pháp như: grey-level quantization, morphology analysis, projection method.

Quá trình nhận dạng (recognition) từng chữ cái có thể gặp nhiều khó khăn vì khi zoom camera có thể dẫn đến kích thước khác nhau của các ký tự. Hơn nữa, nếu ký tự được lấy ra có kích thước quá nhỏ, nó sẽ dễ bị nhầm lẫn về hình dáng với các ký tự khác. Có nhiều phương pháp nhận dạng đã được đề xuất như neural networks, support vector machine (SVM), character templates,...

## II. RELATED WORKS

### A. License plate extraction (LPE)

LPE ảnh hưởng lớn đến độ chính xác đến hệ thống ALPR. Biển số xe có thể được phân biệt thông qua những đặc trưng của nó, có thể là màu, format, những ký tự, kiểu background hay kết hợp các đặc trưng đó lại. Một số phương pháp như:

- Sử dụng đặc trưng viền (boundary): đơn giản nhưng không áp dụng được vào những ảnh phức tạp.
- Sử dụng đặc trưng toàn cục (global): có thể gây ra vật bị vỡ (broken).
- Sử dụng texture: có thể detect được cả khi viền của boundary bị biến dạng, chi phí tính toán cao.
- Sử dụng đặc trưng màu.
- ...

### B. License plate segmentation (LPS)

Ảnh các biển số xe, sau đó sẽ được phân đoạn (segment) để trích xuất ra các ký tự cho phần nhận dạng. Ảnh kết quả từ bước trước có thể có vài vấn đề như bị nghiêng, độ sáng không đồng nhất. Các thuật toán segmentation phải vượt qua được các vấn đề này.

Bilinear transformation được dùng để map ảnh biển số nghiêng thành hình chữ nhật.

Least square để khắc phục ảnh biển số bị nghiêng ngang và nghiêng đứng.

Karhunen-Loeve cũng được sử dụng để khắc phục vấn đề nghiêng của biển số.

Phương pháp line fitting dựa trên bình phương tối thiểu kết hợp với độ lệch vuông góc để chỉnh nghiêng ngang. Để chỉnh nghiêng đứng, ta sẽ cực tiểu hóa phương sai tọa độ của các điểm chiếu được đề xuất.

Chọn threshold không phù hợp cho ảnh biển số nhị phân có thể gây ra việc chồng lên nhau của các ký tự. Điều này làm cho việc segment các ký tự trở nên rất khó. Nâng cao chất

lượng ảnh trước khi nhị phân hóa (binarization) sẽ giúp ta chọn được một threshold phù hợp. Một số kỹ thuật để nâng cao chất lượng ảnh là histogram equalization, contrast enhancement. Một số giải pháp:

- Sử dụng pixel connectivity: phân đoạn bằng cách gán nhãn các pixel có kết nối với nhau trong ảnh nhị phân, đơn giản nhưng thất bại khi những ký tự bị chồng lên nhau hay bị vỡ (broken).
- Sử dụng thông tin chiều: chiếu ảnh nhị phân lên trục thẳng đứng để xác định điểm bắt đầu và kết thúc của ký tự và sau đó chiếu lên trục ngang để lấy từng ký tự. Ưu điểm là lấy ra các ký tự độc lập với vị trí của chúng kể cả khi biển số xe có bị xoay một chút nhưng phụ thuộc nhiều vào chất lượng ảnh.
- Sử dụng những tri thức biết trước (prior knowledge) về các ký tự: bất cứ sự thay đổi nào của các ký tự cũng dẫn đến lỗi.
- Sử dụng kết hợp các đặc trưng: có độ tin cậy cao, nhưng độ tính toán phức tạp.

### C. Character recognition (CR)

Những ký tự được trích xuất ra sẽ được nhận dạng và output trả về là các ký tự trên biển số xe. CR trong ALPR có thể gặp một số khó khăn. Do mức zoom của camera, các ký tự được lấy ra có thể không cùng kích thước hay không cùng mức độ dày nét. Resize kích thước các ký tự lại cùng một cỡ giúp vượt qua được vấn đề này. Ngoài ra, ký tự được lấy ra có thể bị nhiễu (noise), vỡ (broken) hay bị nghiêng (tilted).

Một số giải pháp:

- Sử dụng giá trị các pixel: đơn giản nhưng tốn thời gian.
- Sử dụng các feature rút trích được: có khả năng phát hiện các đặc trưng nổi bật, mạnh khi chống lại các distortion, nhận diện nhanh vì số lượng feature nhỏ hơn số lượng các pixel. Tuy nhiên, quá trình rút trích đặc trưng cần nhiều thời gian.

## III. MÔ HÌNH HỆ THỐNG (SYSTEM MODEL)

Hệ thống ALPR đề xuất cho việc quản lý và điều khiển lưu thông trong các bãi đỗ xe của các tổ chức cá nhân hoặc công cộng thông qua việc định danh biển số xe (vehicle license plate numbers) tại cổng gửi xe. Hệ thống này còn giúp phát hiện trộm cắp xe hoặc theo vết đối tượng khả nghi đang lưu thông bằng phương tiện giao thông trên đường. Hệ thống này chỉ yêu cầu cài đặt camera ở cổng vào/ra tại bãi đỗ. Các hình ảnh thu được từ camera sẽ được xử lý trong máy tính. Tất cả thông tin về phương tiện lưu thông (gồm phần ảnh xe có chứa biển số) sẽ được lưu trữ trong một khoảng thời gian dài. Do vậy, các thông tin chi tiết sẽ được rút trích khi cần thiết.

Hệ thống ALPR rút trích số biển số xe (license plate numbers) từ ảnh đầu vào. Một hệ thống gồm 4 giai đoạn chính. Giai đoạn một thực hiện thu hình ảnh của xe bằng camera. Các thông số của camera như loại camera, độ phân giải của camera (camera resolution), tốc độ màn trập (shutter speed) và điều kiện sáng là các yếu tố cần phải xem xét. Giai đoạn hai thực hiện trích xuất biển số xe từ ảnh dựa vào các đặc trưng như đường biên (boundary), màu và các ký tự. Giai đoạn

ba thực hiện phân vùng biển số và rút trích các ký tự bằng cách chiếu thông tin về màu của chúng, thực hiện gán nhãn hoặc so khớp vị trí của nó trong mẫu. Và giai đoạn cuối thực hiện nhận dạng ký tự (Optical character recognition) sử dụng bộ phân lớp (classifier) như mạng nơron (neural networks).



Fig. 1: Bốn giai đoạn của hệ thống ALPR

Các phần còn lại được tổ chức như sau. Trong phần IV, thu nhận hình ảnh xe (Image acquisition) giới thiệu sơ lược về máy ảnh, định luật khúc xạ ánh sáng và cũng như cách thức giúp đạt được ảnh chất lượng. Phần V, trích xuất vùng có chứa biển số xe (license plate extraction) đưa ra cách thức phân lớp với cái nhìn chi tiết. Phần VI đưa ra phương pháp phân vùng ký tự trong biển số (License plate segmentation) và phần VII thảo luận về cách thức nhận dạng ký tự (OCR).

Ở mỗi phần, chúng tôi xác định bài toán, đưa ra giải pháp hoặc vận dụng thuật toán, thư viện có sẵn để giải quyết vấn đề dựa trên sự tìm hiểu của chúng tôi. Đồng thời phần VIII, chúng tôi tóm lược lại bài báo này và đưa ra các mặt cần cải thiện cho tương lai của nghiên cứu.

## IV. THU NHẬN HÌNH ẢNH XE (IMAGE ACQUISITION)

Kết quả tốt nhất có thể đạt được là với ảnh hồng ngoại (infrared - IR) bởi vì các bước phân vùng (segmentation) giúp cho phát hiện và OCR trở nên dễ hơn, sạch sẽ và giảm thiểu lỗi. Điều này là do các định luật phản xạ ánh sáng, cơ bản là góc tới (angle of incidence) sẽ bằng góc phản xạ (angle of reflection); chúng ta có thể nhìn thấy định luật cơ bản này khi chúng đi qua bề mặt nhẵn như gương phẳng. Sự phản chiếu của bề mặt nhám như giấy dẫn đến một loại phản xạ gọi là phản xạ khuếch tán (diffuse reflection). Đa số các biển số có một đặc tính gọi là retro-reflection - mặt phẳng của biển số được làm từ một loại vật liệu được bao phủ bởi hàng nghìn bán cầu nhỏ (tiny hemispheres) dẫn đến ánh sáng bị phản xạ ngược lại hướng tới như chúng ta có thể xem ở hình sau đây:

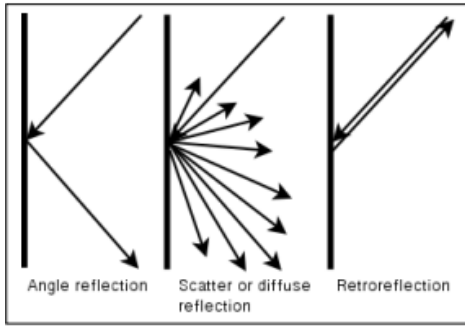


Fig. 2: Các loại phản xạ của ánh sáng

Nếu chúng ta sử dụng camera với bộ lọc kết hợp với máy chiếu sáng hồng ngoại có cấu trúc, chúng ta có thể thu được chỉ ánh sáng hồng ngoại và có được ảnh chất lượng cao (high-quality) để phân vùng, phát hiện và nhận diện số biển số độc lập với mọi môi trường sáng.



Fig. 3: Segmentation

Nhưng trong phạm vi môn học, chúng tôi lựa chọn dùng camera thường. Vì vậy, chúng tôi không thể đạt được kết quả tốt nhất như mong đợi, thậm chí độ lỗi cho việc phát hiện và nhận diện sẽ có thể cao hơn hoặc thất bại trong một số trường hợp. Tuy nhiên, các bước thực hiện cho cả hai đều như nhau. Chúng tôi làm việc với hình ảnh phía trước xe, chụp trực diện với chiều rộng 800 pixels và được chụp cách xe 2-4 (m). Các yêu cầu này đảm bảo để việc thực hiện phân vùng chính xác hơn. Chúng tôi vẫn có thể thực hiện việc phát hiện biển số nếu chúng tôi tạo ra một thuật toán cho hình ảnh đa tỉ lệ (multi-scale image).

## V. LICENSE PLATE EXTRACTION

Giai đoạn này ảnh hưởng đáng kể đến độ chính xác của hệ thống ALPR, đây có thể gọi là bước tiền xử lý (preprocess). Input là ảnh chụp trực diện phần trước của xe (xe oto) và Output là một phần hay một vùng tiềm năng của ảnh chứa biển số. Biển số có thể nằm ở bất kỳ đâu trong ảnh. Thay vì phải duyệt từng pixel trong ảnh, tốn thời gian, ta có thể dùng các đặc trưng (features) để phân biệt biển số và do vậy giảm thời gian xử lý cho hệ thống. Các đặc trưng ấy được lấy từ định dạng và các kí tự cấu thành biển số.

Các đặc trưng có thể là màu sắc, đường biên (boundary) của khung biển số. Sự biến đổi màu sắc giữa các kí tự và nền biển số được gọi là kết cấu (texture) cũng được xem là đặc trưng. Sự tồn tại của các kí tự được dùng như đặc trưng để xác định vùng của biển số. Các đặc trưng có thể kết hợp với nhau để định vị vùng biển số (license plate localization).

Trong phạm vi nghiên cứu, chúng tôi lựa chọn đường biên hoặc biên cạnh như các đặc trưng để định vị vùng chứa biển

số. Vì các biển số thông thường đều có dạng hình chữ nhật với một tỉ lệ khung hình biết trước (tùy vào từng quốc gia), nó có thể được trích xuất bởi tất cả các phần vùng hình chữ nhật khả dĩ có trong ảnh. Phát hiện biên cạnh (Edge detection) thường được sử dụng để tìm các hình chữ nhật đó.

Chúng tôi chia vấn đề này thành hai bước xử lý: phân chia vùng quan tâm (region of interest segmentation) và phân lớp (classification). Bước đầu tiên (region of interest segmentation), chúng tôi áp dụng các bộ lọc khác nhau (different filters), các phép tính hình thái học (morphological operations), thuật toán đường viền (contour algorithm), kiểm chứng (validations) để trích xuất các vùng có thể là biển số trong ảnh. Bước hai (classification) chúng tôi ứng dụng bộ phân lớp Support Vector Machine (SVM classifier) cho từng mảng vế của hình ảnh (image patch) thu được từ bước 1 - là các đặc trưng. Chúng ta cần phải huấn luyện trên dữ liệu huấn luyện (training data) với hai lớp khác biệt - là biển số (plate: 1) và không là biển số (non-plate: 0).

Chúng tôi liệt kê các bước liên quan đến License plate extraction trong hình ảnh dưới đây:

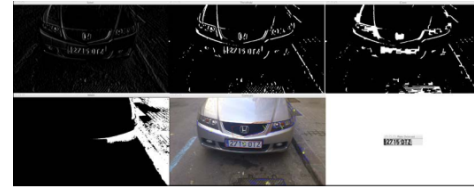


Fig. 4: Qui trình của License plate extraction. Thứ tự từ trái qua và từ trên xuống dưới. Hình (1): Sobel filter. Hình (2): Threshold operation. Hình (3): Close morphologic operation. Hình (4): Mask of one filled area. Hình (5): Các vùng phát hiện có thể chứa biển số được đánh dấu màu đỏ (features images). Hình (6): Ảnh biển số khi dùng SVM classifier.

### A. Region of interest segmentation

Nhiệm vụ là phân chia hình ảnh thành nhiều vùng khác nhau. Qui trình này để đơn giản hóa hình ảnh cho việc phân tích và rút trích đặc trưng dễ dàng hơn. Như hình 4 thì bước này được chia thành 5 bước nhỏ.

Bước 1: Sobel filter được sử dụng để phát hiện biên cạnh. Do sự chuyển màu giữa biển số xe và thân xe, ranh giới của biển số xe được thể hiện bằng các cạnh trong hình ảnh. Một điều quan trọng của Region of interest segmentation là số lượng lớn các cạnh dọc trong biển số được giả định rằng hình ảnh được chụp ở phía trước, và tấm ảnh không được xoay và không bị biến dạng phối cảnh. Đặc trưng này có thể được khai thác trong suốt bước phân đoạn đầu tiên để loại bỏ các khu vực không có bất kỳ cạnh dọc nào. Lưu ý, trước khi tìm cạnh thì ta cần phải chuyển đổi về ảnh grayscale và giảm nhiễu tạo bởi camera hay môi trường xung quanh bằng việc tích chập với Gaussian kernel  $5 \times 5$ . Sau đó, chúng tôi tìm cạnh dọc bằng tích chập với Sobel filter (đạo hàm theo trục tung).

Bước 2: Tiếp đến, chúng tôi đặt ngưỡng để đạt được ảnh nhị phân (binary image) với giá trị cho ngưỡng dựa vào thuật

toán Otsu. Thuật toán Otsu sẽ tự động tính giá trị ngưỡng tối ưu từ histogram của ảnh.

Bước 3: Tiếp theo, để loại bỏ các khoảng trống giữa các cạnh dọc và kết nối các vùng có nhiều biên cạnh thì ta cần dùng close morphological operation. Trong bước này, chúng ta có các vùng có thể chứa biển số. Tuy nhiên, hầu hết các khu vực sẽ không chứa biển số xe. Các vùng này có thể được phân tách bằng phân tích thành phần kết nối (connected-component analysis). Bước này lấy các đường viền (contour) của ảnh nhị phân với các phương thức và kết quả khác nhau. Đối với mỗi đường viền được phát hiện, trích xuất vùng hình chữ nhật giới hạn của diện tích tối thiểu (the bounding rectangle of minimal area) hay còn gọi là hình chữ nhật xoay (rotated rectangle) vì có thể nó bị nghiêng. Đồng thời, chúng tôi cũng thực hiện kiểm chứng sơ bộ (preliminary validations) trên vùng và tỉ lệ khung hình của nó trước khi dùng SVM để phân lớp. Ở Việt Nam thì biển số xe hơi trước (loại dài) sẽ có chiều cao 11 cm, chiều dài 47 cm. Do đó, tỉ lệ khung hình (aspect ratio) xấp xỉ khoảng  $47/11 = 4.272727$  (width/height) với độ lệch (error margin) 40% và diện tích dao động tối thiểu 15 pixel và tối đa 125 pixel cho chiều cao của biển số. Các giá trị này được tính tùy thuộc vào kích thước hình ảnh và vị trí camera.

Bước 4 (Create mask of filled area): Chúng tôi có cải tiến bằng cách sử dụng tính chất nền trắng của biển số xe. Tất cả các biển số có cùng màu nền và chúng ta có thể sử dụng thuật toán tô màu vết loang (floodfill algorithm) để lấy rotated rectangle với xén ảnh chính xác hơn. Trước tiên để cắt ảnh biển số là đặt một số hạt giống (seeds) gần tâm của rotated rectangle cuối cùng. Sau đó lấy kích thước tối thiểu của biển số nằm giữa chiều rộng và chiều cao, và sử dụng nó để phát sinh hạt giống ngẫu nhiên gần trung tâm bản vá. Chúng ta muốn chọn vùng màu trắng thì cần một số hạt giống để chạm vào ít nhất một pixel trắng. Sau đó, với mỗi hạt giống, chúng tôi sử dụng FloodFill để vẽ mặt nạ mới (new mask) để lưu trữ vùng cắt mới gần nhất. FloodFill sẽ giúp fill các thành phần kết nối (connected component) với màu vào mask image bắt đầu từ hạt giống và đặt độ chênh lệch độ sáng/màu thấp hơn và lớn hơn giữa các pixel để fill và pixel lân cận hoặc các seed pixel. Sau đó, chúng tôi có được rotated rectangle từ image-mask points và kiểm chứng kích thước lại một nữa.

Bước 5: Đây là bước cuối cùng để hoàn thành segmentation. Giờ chúng ta đã có các vùng hợp lệ, chúng ta chỉ việc xén các vùng đó, loại bỏ các phép xoay (nếu có thể), thay đổi kích cỡ ảnh và cân bằng độ sáng các vùng ảnh bị xén bởi vì mỗi ảnh có kích thước khác nhau, điều kiện sáng cũng khác nhau khi bị xén.

## B. Classification

Sau các bước tiền xử lý và phân vùng các phần ảnh khả dĩ, chúng ta cần phải xác định là phân vùng đó có phải là biển số không. Để làm việc này, chúng ta sử dụng Support Vector Machine (SVM).

Support Vector Machine là thuật toán nhận dạng mẫu (pattern recognition) trong nhánh học có giám sát (supervised-learning) và thường được dùng cho bài toán phân lớp nhị phân (binary classification). Supervised learning là học từ việc sử

dụng dữ liệu có dán nhãn (labeled data). Ta cần phải huấn luyện mô hình học với một lượng dữ liệu có gán nhãn; nói cách khác, mỗi dữ liệu đầu vào cần phải có một lớp (class).

SVM tạo ra một hoặc nhiều siêu phẳng (hyperplanes) để phân tách các lớp dữ liệu. Xét ví dụ phân lớp cơ bản cho tập dữ liệu 2D gồm hai lớp. Dữ liệu này phải khả tách tuyến tính (linearly separable). SVM tìm kiếm một đường thẳng tối ưu để phân tách dữ liệu, thêm vào đó cực đại hóa lề của một bộ phân lớp tuyến tính (linear classifier) là chiều rộng mà ranh giới có thể được tăng lên trước khi gặp phải một điểm dữ liệu, được gọi là hard margin classification. Tức là các điểm dữ liệu đều phân lớp đúng.

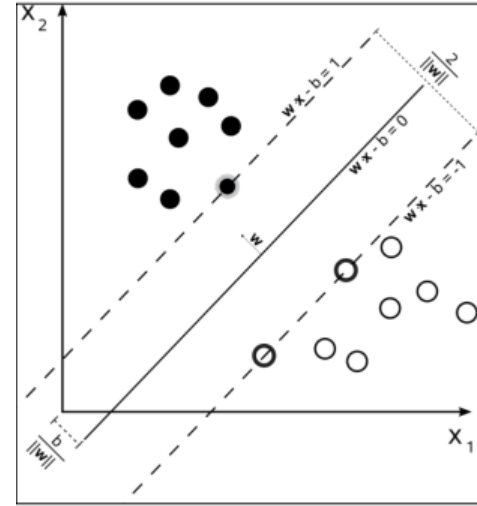


Fig. 5: Linear SVM (Hard margin classification)

Chúng ta có:

$$\begin{cases} y_i = +1, w'x_i + b \geq 1 \\ y_i = -1, w'x_i + b \leq -1 \end{cases}$$

Suy ra, cực đại hóa margin:  $M = \frac{2}{\|w\|}$  tương tự với cực tiểu hóa  $\frac{1}{2}\|w\|^2$  sao cho  $y_i(w'x_i + b) \geq 1 \forall i$  và  $\|w\|^2 = w'w$ .

Từ các phép toán Lagrangian Dual Problem ta có được support vectors khi  $\alpha_i \neq 0$

$$w = \sum_{i=1}^n \alpha_i y_i x_i = \sum_{i \in SV} \alpha_i y_i x_i$$

lấy  $b$  từ  $y_i(w'x_i + b) - 1 = 0$  với  $x_i$  là support vector

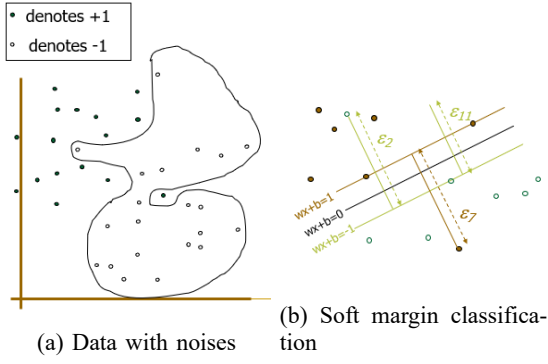
Hàm tách lớp tuyến tính (linear discriminant function):

$$g(x) = w'x + b = \sum_{i \in SV} \alpha_i x'_i x + b \quad (1)$$

Lưu ý:  $x'_i x$  là nhân ma trận (dot product) giữa điểm kiểm tra (test point)  $x$  và support vectors  $x_i$

Nếu dữ liệu có nhiều hay các trường hợp dữ liệu hơi khó để phân tách tuyến tính gọi là slightly non-linear separable thì vẫn có giải pháp là dùng soft margin classification.





Tương tự, ta chỉ cần cực tiểu hóa  $\frac{1}{2}\|w\|^2 + C \sum_{i=1}^n \xi_i$  sao cho  $y_i(wx_i + b) \geq 1 - \xi_i \forall i$  và  $\xi_i \geq 0$ .  $C$  là tham số kiểm soát over-fitting.

Thông thường, nếu dữ liệu không khả tách tuyến tính thì cần phải có biến đổi không tuyến tính (non-linear transformation) để ánh xạ dữ liệu đến không gian có số chiều lớn hơn (higher-dimensional space), nơi mà dữ liệu khả tách, được gọi là không gian đặc trưng (feature space).

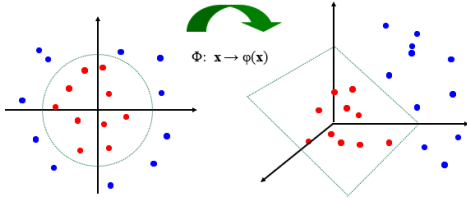


Fig. 7: Non-linear transformation

Không cần phải xác định rõ ràng không gian chiều vì ta chỉ cần nhân hai feature vectors trong cả dữ liệu huấn luyện và dữ liệu kiểm tra. Hàm kernel được xác định như là dot product của 2 feature vectors trong không gian đặc trưng (feature space)

$$K(x_i, x_j) = \phi(x_i)' \phi(x_j)$$

Phương pháp này được gọi là The kernel trick.

- Linear kernel:  $K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j$
- Polynomial kernel:  $K(\mathbf{x}_i, \mathbf{x}_j) = (1 + \mathbf{x}_i^T \mathbf{x}_j)^p$
- Gaussian (Radial-Basis Function (RBF)) kernel:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right)$$

- Sigmoid:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \tanh(\beta_0 \mathbf{x}_i^T \mathbf{x}_j + \beta_1)$$

Fig. 8: Một số hàm kernel thông dụng

Thông qua phép toán Lagrangian dual problem, ta có hàm tách lớp:

$$g(x) = \sum_{i \in SV} \alpha_i K(x_i, x) + b \quad (2)$$

Quay trở lại bài toán ALPR, trước tiên ta cần phải huấn luyện classifier được gọi là offline training. Hạn chế: chúng tôi không có nhiều dữ liệu về biển số nhưng chúng tôi cần khoảng trăm hình ảnh xe hơi để đưa vào tiền xử lý và phân đoạn biển số trong ảnh. Do vậy, chúng tôi sẽ huấn luyện mô hình học với 75 ảnh biển số xe và 35 ảnh không có biển số xe có kích thước 144x33 pixels. Lượng dữ liệu này tuy không nhiều nhưng đủ để đạt được kết quả mong muốn.



Fig. 9: Training examples

Để dễ dàng hiểu cách thức hoạt động của máy học, chúng tôi tạm sử dụng các pixel đặc trưng trong hình ảnh của thuật toán phân lớp. Lưu ý, có các phương pháp và tính năng tốt hơn để đào tạo một SVM, như Phân tích thành phần chính (Principal Components Analysis), biến đổi Fourier (Fourier transform), phân tích kết cấu (texture analysis),... Sau khi huấn luyện với training data, ta có được classifier để dự đoán. Input: Phân ảnh có thể chứa biển số; output: 1 nếu là biển số, 0 nếu không là biển số.

## VI. PHÂN ĐOẠN TỪNG KÝ TỰ TRÊN BIỂN SỐ XE (LICENSE PLATE SEGMENTATION)

Bước cuối kế tiếp là để lấy được các ký tự từ ảnh các biển số bằng object character recognition (OCR). Với mỗi ảnh biển số xe, ta phân đoạn (segment) biển đó riêng ra từng ký tự.

a) *OCR segmentation*: Đầu tiên, sử dụng Histogram Equalization trên ảnh biển số xe, sau đó áp một threshold filter lên ảnh đó để chỉ giữ lại những ký tự. Ảnh threshold có được sẽ được dùng để tìm contour của các ký tự. Mỗi contour phát hiện được (tương ứng một ký tự) cần phải được xác minh lại về kích thước (size verification) để loại bỏ những vùng kích thước quá nhỏ hay tỉ lệ aspect (aspect ratio) chưa đúng lắm, vẫn nên có một khoảng phần trăm sai số cho các ký tự bị rotated hay distorted. Để đếm diện tích, ta sẽ đếm số lượng các non-zero pixel.

Ký tự nào đã được verified thì ta phải tiền xử lý nó về đúng kích thước - vị trí và lưu nó vào một vector các ký tự có thứ tự lấy ra từ ảnh biển số.

## VII. NHẬN DẠNG KÝ TỰ (CHARACTER RECOGNITION)

Cuối cùng, sử dụng Artificial Neural Network (ANN) để nhận dạng ký tự.

a) *Feature extraction*: Trước tiên, sẽ phải rút trích đặc trưng cho từng ký tự vừa được lấy ra. Ở đây, ta sẽ sử dụng ảnh low-resolution và histogram tích lũy (accumulation) theo chiều dọc và ngang cho từng ký tự. Các histogram này cũng cần được chuẩn hóa. Để chọn được resolution thích hợp (5x5

hay 10x10 hay 20x20, ...) thì cần đánh giá thử với từng kích thước, cái nào trả về kết quả tốt nhất sẽ được chọn trong hệ thống.

b) *OCR classification*: Ở bước phân loại, ta sẽ sử dụng Artificial Neural Network (ANN). Cụ thể hơn, là **Multi-Layer Perceptron** (MLP), một trong những thuật toán ANN được dùng nhiều nhất hiện nay. MLP bao gồm mạng các neuron với input layer, output layer và một hay nhiều hidden layer. Mỗi lớp sẽ có ít nhất một neuron sẽ liên kết được với neuron trong lớp trước và neuron trong lớp sau. Mỗi neuron có vài liên kết input vào và vài liên kết output ra. Mỗi neuron tính toán giá trị output bằng cách tính tổng có trọng số các giá trị input cộng với bias term và được chuyển đổi bằng activation function được chọn.

Một trong số các activation function sau đây sẽ được sử dụng: Identity, Sigmoid và Gaussian. Thường được dùng nhất là hàm Sigmoid.

Weight của mỗi layer, synapses và neuron được tính bằng cách train mạng ANN.

Chúng ta cần phải xác định trước số lượng hidden layer cũng như số lượng neuron cần thiết trong mỗi layer để khởi tạo mạng ANN. Kết quả trả về của ANN sẽ là một vector với kích thước là số lượng các class và giá trị ở từng phần tử là xác suất thuộc về class đó.

## VIII. TÓM LƯỢC VÀ CẢI TIẾN

### A. Tóm lược

Nói chung, một hệ thống ALPR bao gồm bốn xử lý các giai đoạn. Trong giai đoạn thu nhận hình ảnh, một số điểm phải được cân nhắc khi chọn camera, chẳng hạn như độ phân giải camera và tốc độ màn trập. Trong giai đoạn trích xuất biển số, biển số được trích xuất dựa trên một số tính năng như màu sắc, ranh giới hoặc dựa trên các ký tự. Trong giai đoạn phân vùng biển số xe, các ký tự được trích xuất bằng cách chiếu thông tin màu của chúng. Cuối cùng, các ký tự nhận dạng trong giai đoạn nhận dạng ký tự bằng mạng lưới neuron (neuron network). ALPR khá khó khăn do sự khác biệt định dạng biển số xe và điều kiện môi trường khác nhau. Có rất nhiều kỹ thuật ALPR đã được đề xuất trong những năm gần đây. Tuy nhiên, vẫn còn thiếu cách thức thống nhất để đánh giá một phương pháp thể nào là hiệu quả. Dĩ nhiên, giải pháp mà chúng tôi đề xuất ở đây vẫn chưa phải là tốt nhất nhưng chúng tôi tiếp tục cải thiện với khả năng mà chúng tôi có.

### B. Cải tiến

Hiện tại, hệ thống ALPR của chúng tôi chỉ xử lý đối với ảnh biển trước xe hơi (loại dài). Chúng tôi, sẽ tiếp tục cải tiến và ứng dụng cho biển sau của xe hơi và cả xe moto.

Chúng tôi vẫn chưa xử lý được ảnh đầu vào bị xoay hay nghiêng. Chúng tôi sẽ tiến hành cải tiến để hệ thống có thể kháng với phép xoay. Kể cả ảnh chất lượng thấp hoặc ảnh bị xuyên tạc.

Trong giai đoạn hai (phần V), chúng tôi chỉ dùng đặc trưng cạnh để phát hiện các vùng có thể là biển số. Chúng tôi sẽ tiếp tục tìm hiểu về các phương pháp khác như: Texture features,

Color features hoặc kết hợp hai hay nhiều loại đặc trưng... Bên cạnh đó, nếu được chúng tôi sẽ dùng thêm PCA hoặc Fourier transform để hỗ trợ cho việc phân lớp vùng ảnh có phải là biển số không.

## REFERENCES

- [1] Baggio, D. L. (2012). 5. Number Plate Recognition Using SVM and Neural Networks. In *Mastering OpenCV with Practical Computer Vision Projects* (6th ed., pp. 161-188). Birmingham, UK: Packt Publishing.
- [2] Du, S., Ibrahim, M., Shehata, M., & Badawy, W. (2013). Automatic license plate recognition (ALPR): A state-of-the-art review. *IEEE Transactions on circuits and systems for video technology*, 23(2), 311-325.
- [3] Kumari, S., Gupta, L., & Gupta, P. (2017). Automatic License Plate Recognition Using OpenCV and Neural Network. *International Journal of Computer Science Trends and Technology (IJCTST)*, 5(3), 114-118.
- [4] Saghaei, H. (2016). Proposal for Automatic License and Number Plate Recognition System for Vehicle Identification. *arXiv preprint arXiv:1610.03341*.