

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



Nguyễn Thế Quyền

**NGHIÊN CỨU VỀ MẠNG NƠON TÍCH CHẬP VÀ ỨNG DỤNG
CHO BÀI TOÁN NHẬN DẠNG BIỂN SỐ XE**

LUẬN VĂN THẠC SĨ KỸ THUẬT
(Theo định hướng ứng dụng)

HÀ NỘI - NĂM 2020

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



Nguyễn Thế Quyền

**NGHIÊN CỨU VỀ MẠNG NƠON TÍCH CHẬP VÀ ỨNG DỤNG
CHO BÀI TOÁN NHẬN DẠNG BIỂN SỐ XE**

Chuyên ngành : KHOA HỌC MÁY TÍNH

Mã số : 8.48.01.01

LUẬN VĂN THẠC SĨ KỸ THUẬT

(Theo định hướng ứng dụng)

NGƯỜI HƯỚNG DẪN KHOA HỌC: TS. PHẠM HOÀNG DUY

HÀ NỘI – NĂM 2020

LỜI CAM ĐOAN

Tôi cam đoan đây là công trình nghiên cứu của riêng tôi. Nội dung của luận văn có tham khảo và sử dụng các tài liệu, thông tin được đăng tải trên những tạp chí và các trang web theo danh mục tài liệu tham khảo. Tất cả các tài liệu tham khảo đều có xuất xứ rõ ràng và được trích dẫn hợp pháp.

Tôi xin hoàn toàn chịu trách nhiệm và chịu mọi hình thức kỷ luật theo quy định cho lời cam đoan của mình.

Hà nội, ngày tháng năm 2020

Người cam đoan

Nguyễn Thế Quyền

LỜI CẢM ƠN

Trong quá trình thực hiện luận văn này, Học viên luôn nhận được sự hướng dẫn, chỉ bảo rất tận tình của Thầy **TS. Phạm Hoàng Duy**, giảng viên Khoa Công nghệ Thông tin 1 là cán bộ trực tiếp hướng dẫn khoa học. Thầy đã dành nhiều thời gian trong việc hướng dẫn học viên cách đọc tài liệu, thu thập và đánh giá thông tin cùng phương pháp nghiên cứu để hoàn thành một luận văn cao học.

Học viên xin chân thành cảm ơn các Thầy, Cô giáo trong Học viện Công nghệ Bru chính Viễn thông đã luôn nhiệt tình giúp đỡ và tạo điều kiện tốt nhất cho em trong suốt quá trình học tập tại trường.

Xin chân thành cảm ơn các anh, các chị và các bạn học viên lớp Cao học – trong Học viện đã luôn động viên, giúp đỡ và nhiệt tình chia sẻ với em những kinh nghiệm học tập, công tác trong suốt khoá học.

Học viên cũng xin chân thành cảm ơn các vị lãnh đạo và các bạn đồng nghiệp tại cơ quan đã luôn tạo mọi điều kiện tốt nhất để em có thể hoàn thành tốt đẹp khoá học Cao học này.

Em xin chân thành cảm ơn !

Hà Nội, ngày tháng năm 2020

MỤC LỤC

GIẢI THÍCH ĐỊNH NGHĨA VÀ TỪ VIẾT TẮT.....	v
DANH SÁCH HÌNH	vi
DANH SÁCH BẢNG	vii
PHẦN MỞ ĐẦU.....	1
CHƯƠNG 1. MẠNG NƠRON VÀ MẠNG NƠRON TÍCH CHẬP.....	2
1.1. Giới thiệu về mạng nơron	2
1.2. Mạng nơron tích chập	5
1.2.1. Định nghĩa mạng nơron tích chập	5
1.2.2. Convolution (tích chập)	5
1.3. Mô hình mạng nơron tích chập	5
1.4. Tính chất của mạng nơron tích chập	8
1.5. Kết luận chương.....	10
CHƯƠNG 2. TỔNG QUAN VỀ NHẬN DẠNG BIỂN SỐ XE.....	11
2.1. Khái niệm về hệ thống nhận dạng biển số xe.....	11
2.1.1. Lịch sử phát triển	11
2.1.2. Cách thức hoạt động của hệ thống nhận dạng biển số xe	11
2.1.3. Phân loại các ứng dụng nhận dạng biển số xe	12
2.1.4. Ứng dụng thực tiễn tại Việt Nam.....	13
2.1.5. Phân loại biển số xe	14
2.2. Phương pháp nhận dạng biển số xe từ ảnh chụp camera.....	17
2.2.1. Phương pháp chuyển đổi Hough.....	17
2.2.2. Phương pháp hình thái học.....	18
2.3. Phương pháp nhận dạng ký tự trong biển số xe.....	19
2.4. Phạm vi nghiên cứu và hướng giải quyết	20
2.5. Kết luận chương.....	22

CHƯƠNG 3. ỨNG DỤNG MẠNG NƠON TÍCH CHẬP TRONG NHẬN	
DẠNG KÝ TỰ	23
3.1. Nhận dạng ký tự trong bài toán nhận dạng biển số xe	23
3.2. Giai đoạn xử lý ảnh	23
3.2.1. Cách phân loại ảnh	23
3.2.2. Các giai đoạn xử lý ảnh.....	24
3.2.3. Xác định vùng chứa biển số	25
3.3. Xây dựng mô hình nhận dạng ký tự	28
3.4. Thuật toán	31
3.4.1. Tách dòng.....	31
3.4.2. Tách ký tự	32
3.4.3. Ánh xạ vào ma trận	32
3.4.4. Huấn luyện mạng nơon.....	34
3.4.5. Nhận dạng ký tự	34
3.5. Cài đặt ứng dụng	35
3.5.1. Môi trường cài đặt và các yêu cầu phần cứng phần mềm.....	35
3.5.2. Giao diện chương trình chính	36
3.6. Kết quả thực nghiệm nhận dạng biển số xe	37
3.7. Kết luận chương	38
KẾT LUẬN	40
DANH MỤC TÀI LIỆU THAM KHẢO	41

GIẢI THÍCH ĐỊNH NGHĨA VÀ TỪ VIẾT TẮT

Định nghĩa/ Từ viết tắt	Giải thích
CNTT	Công nghệ thông tin
CSDL	Cơ sở dữ liệu
ND	Người dùng
MLP	Thuật toán lan truyền ngược
RGB	Red Green Blue
ANPR	Automatic number Plate Recognition

DANH SÁCH HÌNH

Hình 1-1. Cấu tạo một Noron.....	2
Hình 1-2. Minh họa tích chập	5
Hình 1-3. Mô hình mạng noron tích chập.....	6
Hình 1-4. Tích chập trên bản đồ kích hoạt.....	8
Hình 1-5. Kết nối cục bộ.....	9
Hình 1-6. Quá trình tổng hợp	9
Hình 2-1. Ảnh xám và lược đồ xám của ảnh	19
Hình 3-1. Ảnh màu, ảnh mức xám và ảnh nhị phân	24
Hình 3-2. Các bước xử lý ảnh cơ bản	24
Hình 3-3. Một số loại biển số xe thông dụng.....	26
Hình 3-4. Các bước phát hiện vùng chứa biển số xe	27
Hình 3-5. Bộ lọc được sử dụng trong lớp tích chập đầu tiên là các ma trận kích thước 3x3 của -1,0 và 1	29
Hình 3-6. Nhân chập bộ lọc F_1 với ma trận ảnh đầu vào của số 7	30
Hình 3-7. Ví dụ về bộ lọc cạnh (đứng phải, đứng trái, ngang dưới, ngang trên) với đầu vào là ảnh số viết tay.	31
Hình 3-8. Quá trình chia lưới kí tự.....	33
Hình 3-9. Quá trình ánh xạ từ ma trận điểm sang ma trận giá trị	34
Hình 3-10. Giao diện Visual studio 2015	35
Hình 3-11. Giao diện chương trình chính	36
Hình 3-12. Một số biển không phát hiện được biển số.....	37
Hình 3-13. Một số biển không tách đúng ký tự, chữ số bị dính với các vật bên ngoài như đinh ốc, ký tự bị mờ nét, mất nét, loang lổ,...	37
Hình 3-14. Một số biển số xe nhận dạng sai hặc không thể nhận dạng.....	38
Hình 3-15. Mẫu biển số xe nhận dạng chuẩn.....	38

DANH SÁCH BẢNG

Bảng 1.1. Một số hàm truyền thông dụng	4
Bảng 2.1. Quy định biến số cho quân đội.	15

PHẦN MỞ ĐẦU

Deep Learning là một thuật toán dựa trên một số ý tưởng từ não bộ tới việc tiếp thu nhiều tầng biểu đạt, cả cụ thể lẫn trừu tượng, qua đó làm rõ nghĩa của các loại dữ liệu. Deep Learning được ứng dụng trong nhận diện hình ảnh, nhận diện giọng nói, xử lý ngôn ngữ tự nhiên

Hiện nay rất nhiều các bài toán nhận dạng sử dụng deep learning để giải quyết do deep learning có thể giải quyết các bài toán với số lượng lớn, kích thước đầu vào lớn với hiệu năng cũng như độ chính xác vượt trội so với các phương pháp phân lớp truyền thống.

Những năm gần đây, ta đã chứng kiến được nhiều thành tựu vượt bậc trong ngành Thị giác máy tính (Computer Vision). Các hệ thống xử lý ảnh lớn như Facebook, Google hay Amazon đã đưa vào sản phẩm của mình những chức năng thông minh như nhận diện khuôn mặt người dùng, phát triển xe hơi tự lái hay drone giao hàng tự động.

Convolutional Nơon Network (CNNs – Mạng nơ-ron tích chập) là một trong những mô hình Deep Learning tiên tiến giúp cho chúng ta xây dựng được những hệ thống thông minh với độ chính xác cao như hiện nay. Trong luận văn cao học này, em đi vào nghiên cứu về mạng nơron cũng như mạng Convolution (tích chập) cũng như ý tưởng của mô hình CNNs trong phân lớp ảnh (Image Classification), và áp dụng trong việc xây dựng hệ thống nhận dạng biển số xe tự động.

Nội dung bài báo cáo bao gồm 3 chương.

Chương 1: Mạng nơron và mạng nơron tích chập.

Chương 2: Tổng quan về nhận dạng biển số xe.

Chương 3: Áp dụng mạng nơron tích chập trong nhận dạng ký tự

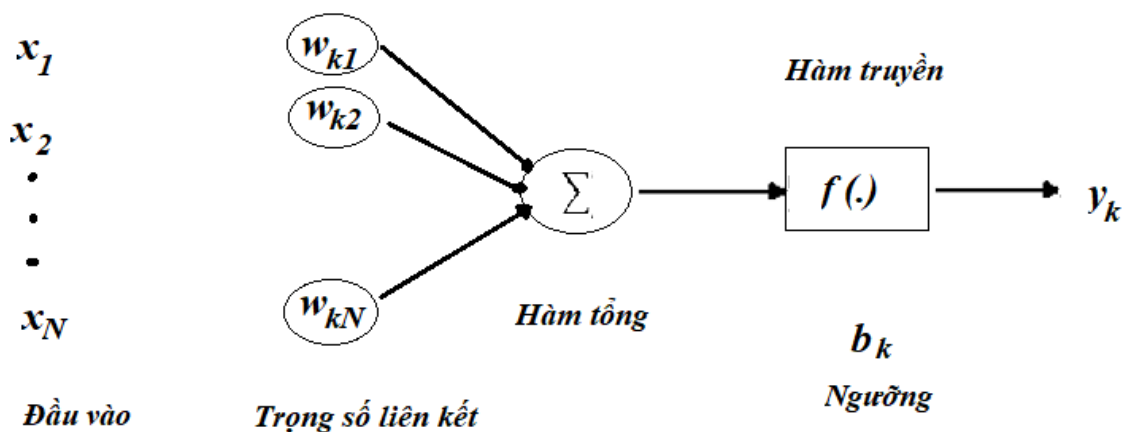
Phần kết luận tóm tắt lại các nội dung đã đạt được của luận văn, và nêu lên một số gợi ý về hướng phát triển tiếp theo của luận văn.

CHƯƠNG 1. MẠNG NƠRON VÀ MẠNG NƠRON TÍCH CHẬP

1.1. Giới thiệu về mạng nơron

Định nghĩa: Mạng nơron nhân tạo, Artificial Noron Network (ANN) là một mô hình xử lý thông tin phỏng theo cách thức xử lý thông tin của các hệ nơron sinh học. Nó được tạo nên từ một số lượng lớn các phần tử (nơron) kết nối với nhau thông qua các liên kết (trọng số liên kết) làm việc như một thể thống nhất để giải quyết một vấn đề cụ thể nào đó. Một mạng nơron nhân tạo được cấu hình cho một ứng dụng cụ thể (nhận dạng mẫu, phân loại dữ liệu,...) thông qua một quá trình học từ tập các mẫu huấn luyện. Về bản chất học chính là quá trình hiệu chỉnh trọng số liên kết giữa các nơron.

Cấu trúc nơron nhân tạo:



Hình 1-1. Cấu tạo một Nơron

Các thành phần cơ bản của một nơron nhân tạo bao gồm:

Tập các đầu vào: Là các tín hiệu vào (input signals) của nơron, các tín hiệu này thường được đưa vào dưới dạng một vector N chiều.

Tập các liên kết: Mỗi liên kết được thể hiện bởi một trọng số liên kết – Synaptic weight. Trọng số liên kết giữa tín hiệu vào thứ j với nơron k thường được ký hiệu là w_{kj} . Thông thường, các trọng số này được khởi tạo một cách ngẫu nhiên ở thời điểm khởi tạo mạng và được cập nhật liên tục trong quá trình học mạng.

Bộ tổng (Summing function): Thường dùng để tính tổng của tích các đầu vào với trọng số liên kết của nó.

Ngưỡng (còn gọi là một độ lệch - bias): Ngưỡng này thường được đưa vào như một thành phần của hàm truyền.

Hàm truyền (Transfer function): Hàm này được dùng để giới hạn phạm vi đầu ra của mỗi neuron. Nó nhận đầu vào là kết quả của hàm tổng và ngưỡng.

Đầu ra: Là tín hiệu đầu ra của một neuron, với mỗi neuron sẽ có tối đa là một đầu ra.

Xét về mặt toán học, cấu trúc của một neuron k , được mô tả bằng cặp biểu thức sau:

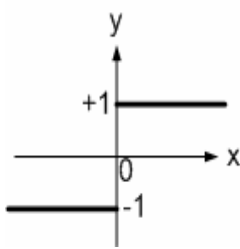
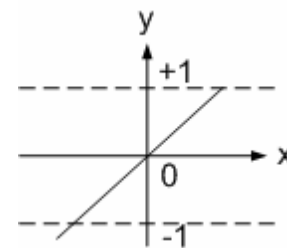
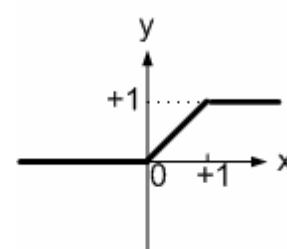
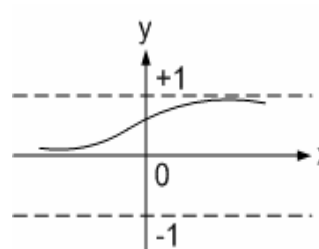
$$u_k = \sum_{j=1}^p w_{kj} x_j$$

Và $y_k = f(b_1 - b_2)$

Trong đó: x_1, x_2, \dots, x_p : là các tín hiệu vào; $(w_{k1}, w_{k2}, \dots, w_{kp})$ là các trọng số liên kết của neuron thứ k ; u_k là hàm tổng; b_k là một ngưỡng; f là hàm truyền và y_k là tín hiệu đầu ra của neuron.

Như vậy neuron nhân tạo nhận các tín hiệu đầu vào, xử lý (nhân các tín hiệu này với trọng số liên kết, tính tổng các tích thu được rồi gửi kết quả tới hàm truyền), và cho một tín hiệu đầu ra (là kết quả của hàm truyền).

Bảng 1.1. Một số hàm truyền thông dụng

Hàm truyền	Đồ thị	Định nghĩa
Symmetrical Hard Limit (hardlims)		Hàm này cũng được biết đến với tên "Hàm ngưỡng" (Threshold function hay Heaviside function). Đầu ra của hàm này được giới hạn vào một trong hai giá trị: $g(x) = \begin{cases} 1, & \text{nếu } (x \geq \theta) \\ 0, & \text{nếu } (x < \theta) \end{cases}$
Linear (purelin)		Nếu coi các đầu vào là một đơn vị thì chúng sẽ sử dụng hàm này. Đôi khi một hằng số được nhân với net-input để tạo ra một hàm đồng nhất
Saturating Linear (satlin)		Hàm này có các thuộc tính tương tự hàm sigmoid. Nó làm việc tốt đối với các ứng dụng có đầu ra yêu cầu trong khoảng $[-1,1]$.
Log-Sigmoid (logsig)		Hàm này đặc biệt thuận lợi khi sử dụng cho các mạng được huấn luyện (trained) bởi thuật toán Lan truyền ngược (back-propagation), bởi vì nó dễ lấy đạo hàm, do đó có thể giảm đáng kể tính toán trong quá trình huấn luyện. Hàm này được ứng dụng cho các chương trình ứng dụng mà các đầu ra mong muốn rơi vào khoảng $[0,1]$.

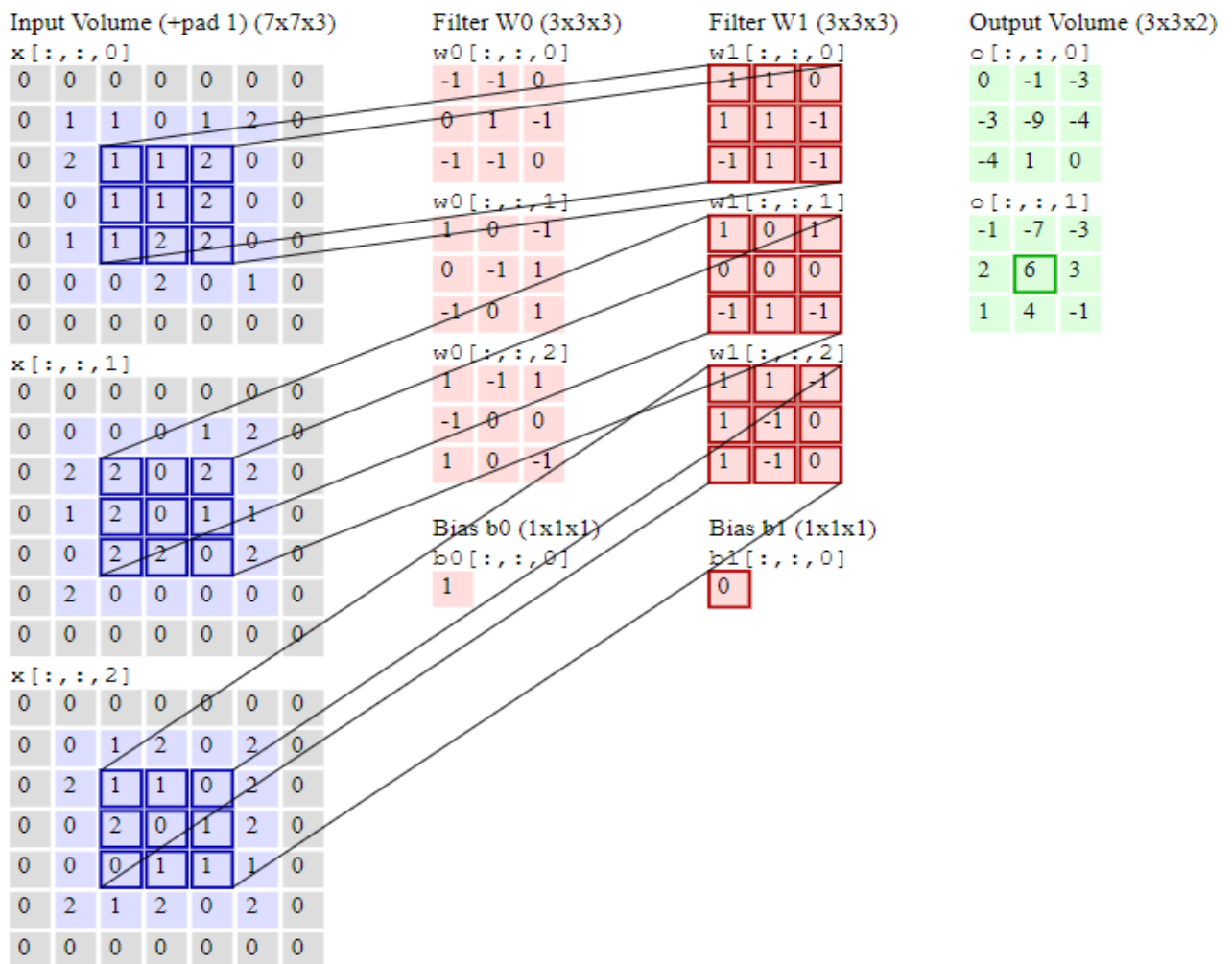
1.1. Mạng nơ-ron tích chập

1.1.1. Định nghĩa mạng nơ-ron tích chập

Tích chập được ứng dụng phổ biến trong lĩnh vực thị giác máy tính. Thông qua các phép tích chập, các đặc trưng chính từ ảnh được chiết xuất và truyền vào các lớp *tích chập* (layer convolution). Mỗi một lớp tích chập sẽ bao gồm nhiều đơn vị mà kết quả ở mỗi đơn vị là một phép biến đổi tích chập từ layer trước đó thông qua phép nhân tích chập với bộ lọc.

1.1.2. Convolution (tích chập)

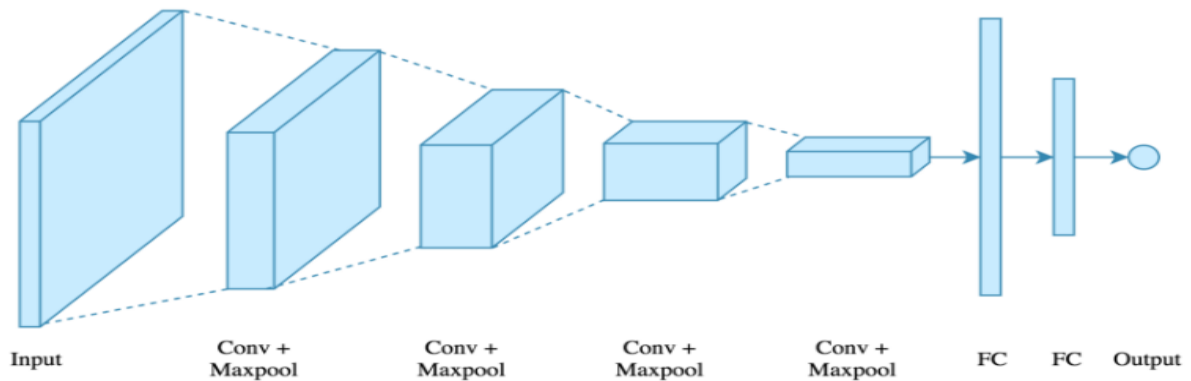
Tích chập là một khái niệm trong xử lý tín hiệu số nhằm biến đổi thông tin đầu vào thông qua một phép tích chập với bộ lọc để trả về đầu ra là một tín hiệu mới. Tín hiệu này sẽ làm giảm những đặc trưng mà bộ lọc không quan tâm và chỉ giữ những đặc trưng chính.



Hình 1-2. Minh họa tích chập

Ảnh đầu vào được cho qua một bộ lọc chạy dọc bức ảnh. Bộ lọc có kích thước là (3x3 hoặc 5x5) và áp dụng phép tích vô hướng để tính toán, cho ra một giá trị duy nhất. Đầu ra của phép tích chập là một tập các giá trị ảnh được gọi là mạng đặc trưng (features map).

1.2. Mô hình mạng nơron tích chập



Hình 1-3. Mô hình mạng nơron tích chập

CNNs có tính bất biến và tính kết hợp cục bộ (Location Invariance and Compositionality). Với cùng một đối tượng, nếu đối tượng này được chiếu theo các góc độ khác nhau (translation, rotation, scaling) thì độ chính xác của thuật toán sẽ bị ảnh hưởng đáng kể. Pooling lớp sẽ cho bạn tính bất biến đối với phép dịch chuyển (translation), phép quay (rotation) và phép co giãn (scaling).

Về cơ bản thiết kế của một mạng nơron tích chập 2 chiều có dạng như sau:

INPUT -> [[CONV -> RELU] N -> POOL?] M -> [FC -> RELU] $*K$ -> FC

Trong đó:

- **Input:** Lớp đầu vào
- **Conv Layer (Lớp tích chập):** Mục tiêu của các lớp tích chập là trích chọn các đặc trưng của ảnh đầu vào.
- **ReLU Layer (Lớp biến đổi):** ReLU layer áp dụng các kích hoạt (activation function) $\max(0, x)$ lên đầu ra của Conv Layer, có tác dụng đưa các giá trị âm về thành 0. Layer này không thay đổi kích thước của ảnh và không có thêm bất kì tham số nào. Mục đích của lớp ReLU là đưa ảnh một mức ngưỡng, ở đây là 0. Để loại bỏ các giá trị âm không cần thiết mà có thể sẽ ảnh hưởng cho việc tính toán ở các layer sau đó.

- **Pool layer (Lớp tổng hợp):** Pool Layer thực hiện chức năng làm giảm chiều không gian của đầu và giảm độ phức tạp tính toán của model ngoài ra Pool Layer còn giúp kiểm soát hiện tượng overfitting. Thông thường, Pool layer có nhiều hình thức khác nhau phù hợp cho nhiều bài toán, tuy nhiên Max Pooling là được sử dụng nhiều vào phổ biến hơn cả với ý tưởng cũng rất sát với thực tế con người đó là: Giữ lại chi tiết quan trọng hay hiểu ở trong bài toán này chính giữ lại pixel có giá trị lớn nhất.
- **Fully_Connected (Lớp kết nối hoàn toàn):** Tại lớp mạng này, mỗi một nơ-ron của layer này sẽ liên kết tới mọi nơ-ron của lớp khác. Để đưa ảnh từ các layer trước vào mạng này, buộc phải dãn phẳng bức ảnh ra thành 1 vector thay vì là mảng nhiều chiều như trước. Tại layer cuối cùng sẽ sử dụng 1 hàm kinh điển trong học máy mà bất kì ai cũng từng sử dụng đó là softmax để phân loại đối tượng dựa vào vector đặc trưng đã được tính toán của các lớp trước đó.

Các kí hiệu $[N]$, $[M]$ hoặc $[K]$ ám chỉ cấu trúc bên trong $[]$ có thể lặp lại nhiều lần liên tiếp nhau. M , K là số lần lặp lại. Kí hiệu \rightarrow đại diện cho các lớp liên kế nhau mà lớp đứng trước \rightarrow sẽ làm đầu vào cho lớp đứng sau.

Như vậy ta có thể thấy một mạng nơron tích chập về cơ bản có 3 quá trình khác nhau:

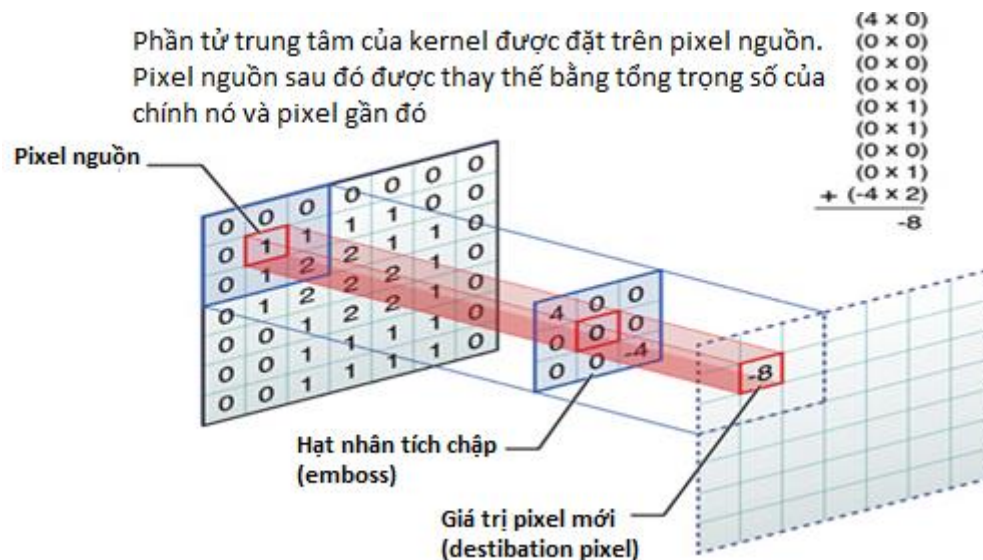
Quá trình chiết xuất đặc trưng: Thông qua các tích chập giữa ma trận đầu vào với bộ lọc để tạo thành các đơn vị trong một lớp mới. Quá trình này có thể diễn ra liên tục ở phần đầu của mạng và thường sử dụng hàm kích hoạt relu.

Quá trình tổng hợp: Các lớp ở về sau quá trình chiết xuất đặc trưng sẽ có kích thước lớn do số đơn vị ở các lớp sau thường tăng tiến theo cấp số nhân. Điều đó làm tăng số lượng hệ số và khối lượng tính toán trong mạng nơron. Do đó để giảm tải tính toán chúng ta sẽ cần giảm chiều của ma trận hoặc giảm số đơn vị của lớp. Vì mỗi một đơn vị sẽ là kết quả đại diện của việc áp dụng 1 bộ lọc để tìm ra một đặc trưng cụ thể nên việc giảm số đơn vị sẽ không khả thi. Giảm kích thước ma trận thông qua việc tìm ra 1 giá trị đại diện cho mỗi một vùng không gian mà bộ lọc đi qua sẽ không làm thay đổi các đường nét chính của bức ảnh nhưng lại giảm được kích thước của ảnh. Do đó quá trình giảm chiều ma trận được áp dụng. Quá trình này gọi là tổng hợp.

Quá trình kết nối hoàn toàn: Sau khi đã giảm số lượng tham số đến một mức độ hợp lý, ma trận cần được làm dẹt (flatten) thành một vector và sử dụng các kết nối hoàn toàn giữa các lớp. Quá trình này sẽ diễn ra cuối mạng tích chập và sử dụng hàm kích hoạt là relu. Kết nối cuối cùng sẽ dẫn tới các đơn vị là đại diện cho mỗi lớp với hàm kích hoạt là softmax nhằm mục đích tính xác suất.

1.3. Tính chất của mạng nơ ron tích chập

Tính kết nối trượt: Khác với các mạng nơ ron thông thường, mạng nơ ron tích chập không kết nối tới toàn bộ hình ảnh mà chỉ kết nối tới từng vùng *vùng địa phương* (local region) có kích thước bằng kích thước bộ lọc của hình ảnh đó. Các bộ lọc sẽ trượt theo chiều của ảnh từ trái qua phải và từ trên xuống dưới đồng thời tính toán các giá trị tích chập và điền vào *bản đồ kích hoạt* (activation map).

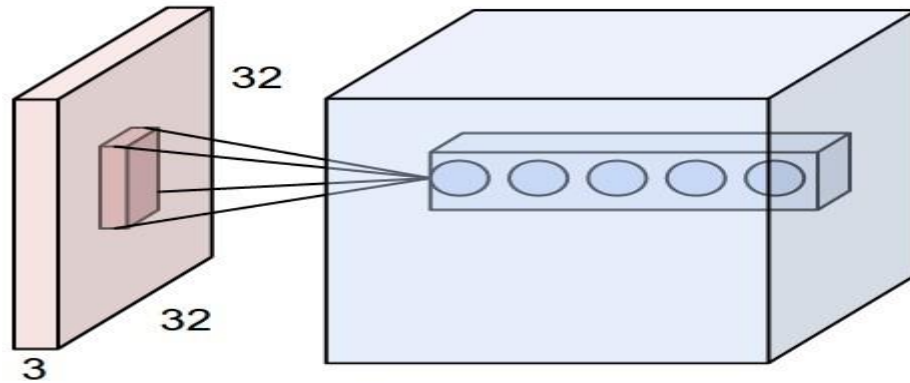


Hình 1-4. Tích chập trên bản đồ kích hoạt

Tính chia sẻ kết nối và kết nối cục bộ: Chúng ta đã biết quá trình biến đổi trong mạng tích chập sẽ kết nối các khối nơ ron 3D. Tuy nhiên các đơn vị sẽ không kết nối tới toàn bộ khối 3D trước đó theo chiều width và height mà chúng sẽ chọn ra các vùng *địa phương* có kích thước bằng với bộ lọc giống như quá trình tính tích chập. Các vùng địa phương sẽ được chia sẻ chung một bộ siêu tham số gọi là trường tiếp nhận (receptive field) của bộ lọc. Tuy nhiên các kết nối cục bộ chỉ diễn ra theo chiều width và height. Kết nối sẽ mở rộng hoàn toàn theo chiều depth. Như vậy số tham số trong một lớp sẽ là $F \times F \times DF \times F \times D$ (F, DF, D lần lượt là kích thước bộ lọc và chiều depth).

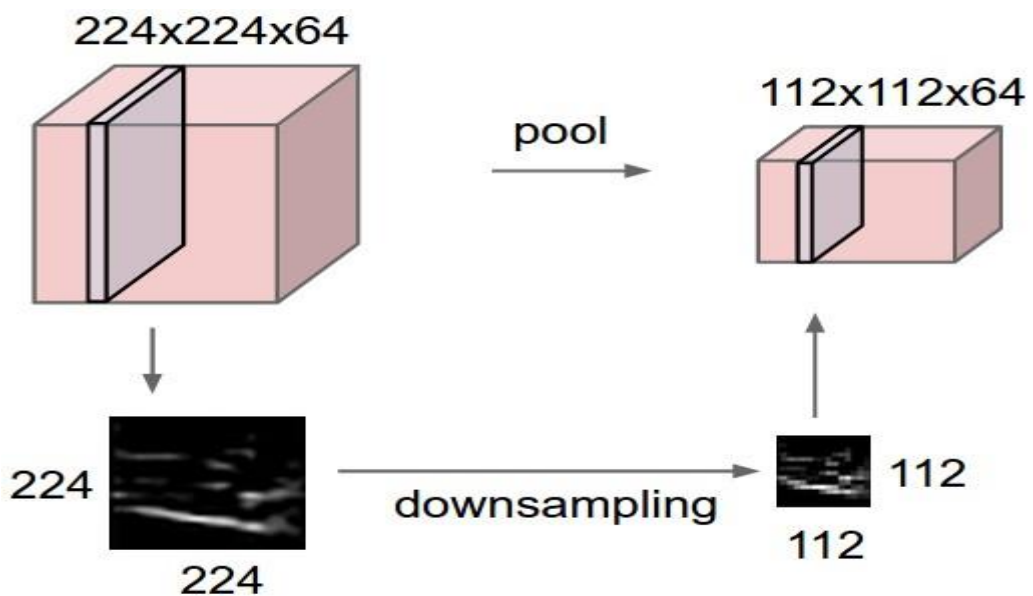
Mỗi bộ lọc đại diện cho một khả năng chiết xuất một đặc trưng nào đó. Do đó khi đi

qua toàn bộ các vùng địa phương của khối nơ ron 3D, các đặc trưng được chiết xuất sẽ hiển thị trên lớp mới.



Hình 1-5. Kết nối cục bộ

Tính tổng hợp: Chúng ta tưởng tượng rằng ở các lớp tích chập gần cuối số tham số sẽ cực kì lớn do sự gia tăng của chiều depth và thông thường sẽ theo cấp số nhân. Như vậy nếu không có một cơ chế kiểm soát sự gia tăng tham số, chi phí tính toán sẽ cực kì lớn và vượt quá khả năng của một số máy tính cấu hình yếu (Như máy của mình chẳng hạn, hơi đáng buồn). Một cách tự nhiên là chúng ta sẽ giảm kích thước các chiều width và height (down sampling) mà vẫn giữ nguyên được các đặc trưng của khối. Các thực hiện tương tự như tính tích chập nhưng thay vì tính tích hadamard giữa ma trận bộ lọc và vùng địa phương ta sẽ tính trung bình (average pooling) hoặc giá trị lớn nhất (max pooling) của các phần tử trong vùng địa phương. Trước đây các tính trung bình được áp dụng nhiều nhưng các mô hình hiện đại đã thay thế bằng giá trị lớn nhất do tốc độ tính max nhanh hơn so với trung bình.



Hình 1-6. Quá trình tổng hợp

Độ phức tạp phát hiện hình ảnh tăng dần: Ở lớp đầu tiên hình ảnh mà chúng ta có chỉ là những giá trị pixels. Sau khi đi qua lớp thứ 2 máy tính sẽ nhận diện được các hình dạng cạnh, rìa và các đường nét đơn giản. Càng ở những lớp tích chập về sau càng có khả năng phát hiện các đường nét phức tạp hoặc vật thể. Đầu ra ở lớp cuối cùng là xác suất thuộc về mỗi lớp.

1.4. Kết luận chương

Rất nhiều các bài toán nhận dạng sử dụng Deep Learning, vì nó có thể giải quyết các bài toán với số lượng lớn các biến, tham số kích thước đầu vào lớn với hiệu năng cũng như độ chính xác vượt trội so với các phương pháp phân lớp truyền thống, xây dựng những hệ thống thông minh với độ chính xác cao.

Convolutional Nơ-ron Network (CNNs – Mạng nơ-ron tích chập) là một trong những mô hình mạng Học sâu phổ biến nhất hiện nay, có khả năng nhận dạng và phân loại hình ảnh với độ chính xác rất cao, thậm chí còn tốt hơn con người trong nhiều trường hợp. Mô hình này đã và đang được phát triển, ứng dụng vào các hệ thống xử lý ảnh lớn của Facebook, Google hay Amazon... cho các mục đích khác nhau, như các thuật toán gắn thẻ tự động, tìm kiếm ảnh hoặc gợi ý sản phẩm cho người tiêu dùng.

CHƯƠNG 2. TỔNG QUAN VỀ NHẬN DẠNG BIỂN SỐ XE

2.1. Khái niệm về hệ thống nhận dạng biển số xe

2.1.1. Lịch sử phát triển

Năm 1992, công nghệ ALPR (Automatic License Plate Number) hay còn gọi là tự động nhận dạng biển số xe, được phát triển tại Đại học Cambridge ở Vương quốc Anh để ứng phó với chủ nghĩa khủng bố.

Đến năm 1996, công nghệ ALPR đã được hoàn thiện tại mỗi công phía tây Vương quốc Anh để đọc tất cả các biển đăng ký xe từ Ireland. Công nghệ ALPR tiếp tục được nghiên cứu và phát triển tại Anh. Kể từ tháng ba năm 2006, hầu hết các con đường, các trung tâm thị trấn, cảng, trạm xăng của London đã được lắp đặt camera chạy phần mềm ALPR.

Trên thế giới hiện nay, bài toán nhận dạng biển số xe được nghiên cứu và phát triển một cách sâu rộng. Nhiều tác giả với các công trình nghiên cứu được công bố với tỉ lệ nhận dạng ngày càng chính xác. Một số bài báo cáo nghiên cứu của các tác giả tiêu biểu trong vài năm trở lại đây như:

Chirag N. Paunwala, 2010 [1] với nội dung: rút trích vùng số xe trong ảnh. Ảnh đầu vào được tiền xử lý bằng cách phương pháp nâng cao chất lượng ảnh, sau đó tìm biên bằng Vertical Edge và xử lý một lần nữa bằng Opening và Closing. Các vùng ứng viên sau đó được kiểm tra bằng thuật toán scan theo dòng để tìm được vùng chứa biển số xe chính xác. Kết quả nhận dạng 750 ảnh trong các điều kiện khác nhau cho tỉ lệ $742/750 = 99.2$.

Choo Kar Soon, 2012 [2] với nội dung: nhận dạng biển số xe tại Malaysia, sử dụng giải thuật Adaboosts để training tập dữ liệu gồm gần 100 ảnh biển số. Các ký tự được nhận dạng bằng phương pháp KNN. Kết quả nhận dạng biển số 98% và nhận dạng ký tự 95% trên ảnh tĩnh.

Báo cáo này nghiên cứu cách nhận dạng biển số xe với sự kết hợp của phép biến đổi Hough và giải thuật tìm Contour để cải thiện kết quả phát hiện. Vùng các ứng viên sau đó tiếp tục được scan theo dòng để đếm số đối tượng bị cắt và so sánh với ngưỡng, nhằm tìm ra vùng ứng viên thỏa mãn. Kết quả nhận dạng đạt 98-99%.

Phần mềm nhận dạng biển số xe, đã được ứng dụng thực tế tại các trạm cân, trạm gửi xe, các trụ đèn giao thông để phát hiện xe vi phạm.

2.1.2. Cách thức hoạt động của hệ thống nhận dạng biển số xe

Trước khi giới thiệu về thư viện Mapkit, tác giả sẽ đưa ra các nội dung cơ bản sau:

Hệ thống ALPR (Automatic License Plate Recognition) gồm phần cứng và phần mềm, trong đó phần cứng là camera thu nhận ảnh xe và phần mềm có chức năng nhận dạng biển số xe từ ảnh chụp của camera. Camera thu nhận ảnh được đặt tại một vị trí cố định sao cho có thể quét được hình ảnh xem một cách rõ ràng và chụp lại hình ảnh đối tượng xe có chứa biển số. Ảnh này được đưa vào phần mềm nhận dạng để trích ra chính xác biển số xe có trong ảnh, sau đó một thuật toán OCR (Optical Character Recognition) được sử dụng để lấy từng ký tự và chuyển đổi thành định dạng mà máy tính có thể phân biệt được các chữ và số như dạng text... Cùng với sự phát triển của công nghệ, camera ngày nay đã có thể chụp một cách rõ nét trong điều kiện xe chạy với tốc độ cao như ở các đường cao tốc.

Không có một hệ thống ALPR nào có thể nhận dạng chính xác 100%. Điều đó phụ thuộc vào nhiều yếu tố như thời tiết, độ sáng, góc của camera tới xe,... Một số yếu tố ảnh hưởng đến độ chính xác của hệ thống là:

- Độ phân giải của ảnh kém hoặc ảnh bị mờ.
- Điều kiện ánh sáng yếu, bị phản chiếu hoặc che bóng.
- Các đối tượng có dạng tương tự như biển số xe ở ngoại cảnh.

Sự khác nhau về cấu trúc biển số xe của mỗi nước.

2.1.3. Phân loại các ứng dụng nhận dạng biển số xe

Có nhiều cách thức khác nhau để phân loại các ứng dụng nhận dạng biển số xe. Một trong những cách đơn giản là phân loại ứng dụng nhận dạng biển số xe thông qua mục đích sử dụng. Có thể chia ứng dụng nhận dạng biển số xe thành hai loại sau:

Loại 1: Giới hạn vùng nhìn

Đầu vào: Ảnh thu trực tiếp từ các thiết bị ghi nhận ảnh kỹ thuật số. Ảnh được ghi nhận thường chỉ giới hạn trong vùng có biển số xe.

Nguyên lý hoạt động: Các phương tiện giao thông phải chạy với một tốc độ đủ chậm để máy ghi nhận hình ảnh có thể thu được ảnh vùng biển số xe.

Ứng dụng: Những ứng dụng nhận dạng biển số xe loại này thường được dùng tại các trạm kiểm soát, các trạm thu phí, các bãi gửi xe tự động, các trạm gác cổng.

Loại 2: Không giới hạn vùng nhìn

Đầu vào: Ảnh đầu vào thu được từ các thiết bị ghi hình tự động, không phụ thuộc vào góc độ, các đối tượng xung quanh, ảnh không cần bắt buộc chỉ chụp vùng chứa biển số xe,

mà có thể ảnh tổng hợp như chứa thêm các đối tượng như người, cây đường phố..., miễn là vùng biên số xe phải đủ rõ để có thể thực hiện nhận dạng được ký tự trong vùng đó.

Nguyên lý hoạt động: Do đặc tính không giới hạn vùng nhìn mà ảnh đầu vào có thể thu được từ một thiết bị ghi hình (camara, máy ảnh...). Và do đó, công việc đầu tiên là dò tìm trong ảnh, để xác định đúng vùng nào là biên số xe. Sau đó, thực hiện tách vùng và nhận dạng. Cuối cùng tùy thuộc vào mục đích sử dụng mà kết quả nhận dạng được truyền đi hay lưu trữ để phục vụ nhu cầu của người dùng cuối.

Ứng dụng: Vì không phụ thuộc vào hình ảnh thu được nên có thể dùng ứng dụng tại nhiều nơi như tại những nơi điều tiết giao thông, tại các vị trí nhạy cảm của giao thông như ngã ba, ngã tư đường giao nhau. Kiểm soát, phát hiện hành vi vi phạm an toàn giao thông.

2.1.4. Ứng dụng thực tiễn tại Việt Nam

Ứng dụng: Vì không phụ thuộc vào hình ảnh thu được nên có thể dùng ứng dụng tại nhiều nơi như tại những nơi điều tiết giao thông, tại các vị trí nhạy cảm của giao thông như ngã ba, ngã tư đường giao nhau. Kiểm soát, phát hiện hành vi vi phạm an toàn giao thông.

Hệ thống nhận dạng biên số xe được xây dựng nhằm mục đích giám sát, kiểm soát các phương tiện. Dưới đây chúng ta đề cập đến một số ứng dụng phổ biến đối với hệ thống nhận dạng biên số xe:

- Thu phí giao thông: Lắp đặt hệ thống “Nhận dạng biên số xe” tại các trạm thu phí nhằm hỗ trợ hoặc tự động hóa công tác thu phí.
- Kiểm soát xe tại các đường biên giới: Mỗi quốc gia đều có những quy định riêng về biên số xe, để phục vụ cho công tác quản lý và phát hiện những phương tiện giao thông (xe) vượt biên giới bất hợp pháp. Việc lắp hệ thống “Nhận dạng biên số xe” tại các trạm kiểm soát sẽ góp phần hỗ trợ công tác kiểm tra và an ninh quốc gia.

- Các trạm gác cổng: Việc lắp đặt hệ thống “Nhận dạng biển số xe” sẽ hỗ trợ hoặc tự động hóa công tác mở cổng cho xe vào. Ngoài ra, hệ thống còn được ứng dụng vào công tác chống trộm xe, các bãi giữ xe tự động, điều tiết giao thông (chẳng hạn như Thành phố Dublin đã ứng dụng công nghệ “Nhận dạng biển số xe tự động” trong việc điều tiết giao thông theo dạng biển số chẵn/lẻ).

2.1.5. Phân loại biển số xe

Quy định về màu sắc và các ký tự trên biển số.

Biển trắng chữ đen dành cho dân sự.

- Màu trắng 2 chữ, 5 số là biển dành cho người nước ngoài.
- NG là xe ngoại giao.
- NN là xe của các tổ chức, cá nhân nước ngoài: Trong đó 3 số ở giữa là mã quốc gia, 2 số tiếp theo là số thứ tự.

(Ghi chú: Xe số 80 NG xxx-yy là biển cấp cho các đại sứ quán, thêm gạch đỏ ở giữa và 2 số cuối là 01 là biển xe của tổng lãnh sự.)

- Biển đỏ chữ trắng là dành cho quân đội.

Bảng 2.1. Quy định biển số cho quân đội.

AT	Binh đoàn 12	AD	QĐ 4, Binh đoàn Cửu Long
BB	Bộ Binh	BC	Binh chủng công binh
BH	Binh chủng hóa học	BS	Binh đoàn Trường Sơn
BT	Binh chủng thông tin liên lạc	BP	Bộ Tư lệnh biên phòng
HB	Học viện lục quân	HH	Học viện quân y
KA	Quân khu 1	KB	Quân khu 2
KC	Quân khu 3	KD	Quân khu 4
KV	Quân khu 5	KP	Quân khu 7
KK	Quân khu 9	PP	Các quân y viện
QH	Quân chủng hải quân	QK,QP	Phòng không không quân
TC	Tổng cục chính trị	QC	Tổng cục hậu cần
TK	Tổng cục CN Quốc phòng	TH	Tổng cục kỹ thuật
TM	Bộ tổng tham mưu	TT	Tổng cục kỹ thuật
		VT	Viettel

Quy định về biển số cho các tỉnh thành:

Bảng 2.2. Quy định biển cho các tỉnh thành

11	Cao Bằng	43	Đà Nẵng	77	Bình Định
12	Lạng Sơn	47	Đắc Lắc	78	Phú Yên
14	Quảng Ninh	48	Đắc Nông	79	Khánh Hòa
15,16	Hải Phòng	49	Lâm Đồng	80	Các đơn vị TW
17	Thái Bình	50-59	TP.HCM	81	Gia Lai
18	Nam định	60	Đồng nai	82	Kon tum
19	Phú thọ	61	Bình dương	83	Sóc trăng
21	Yên Bái	62	Long an	84	Trà vinh
22	Tuyên quang	63	Tiền giang	85	Ninh thuận
23	Hà giang	64	Vĩnh long	86	Bình thuận
24	Lào cai	65	Cần thơ	88	Vĩnh phúc
25	Lai châu	66	Đồng tháp	89	Hưng yên
26	Sơn la	67	An giang	90	Hà nam
27	Điện biên	68	Kiên giang	92	Quảng nam
28	Hòa bình	69	Cà mau	93	Bình phước
29-32	Hà nội	70	Tây Ninh	94	Bạc liêu
33	Hà tây	71	Bến tre	95	Hậu giang
34	Hải dương	72	BR-VT	97	Bắc cạn
35	Ninh bình	73	Quảng bình	98	Bắc giang
36	Thanh hóa	74	Quảng trị	99	Bắc ninh
37	Nghệ an	75	Huế		
38	Hà tĩnh	76	Quảng Ngãi		

Các xe mang biển A: Xe của Công An - Cảnh Sát tương ứng với các tính.

2.2. Phương pháp nhận dạng biển số xe từ ảnh chụp camera

Có nhiều phương pháp để giải quyết vấn đề này nhưng đều quy về các phương pháp chính sau đây:

- Phương pháp dùng chuyển đổi Hough: dựa vào đặc trưng cạnh biên trích được, áp dụng các phương pháp xác định đường thẳng như phép biến đổi Hough để phát hiện các cặp đường thẳng gần song song ghép thành một ảnh biển số. Giao điểm của những đoạn thẳng này chính là vùng bao chứa biển số xe.
- Phương pháp hình thái học: dựa vào đặc trưng hình thái của biển số xe như màu sắc, độ sáng, sự đối xứng... để xác định và trích ra ảnh biển số.
- Phương pháp khớp mẫu: xem biển số là một đối tượng có khung nền riêng và sử dụng các cửa sổ dò để trích từng đối tượng đưa qua mạng nơron (neuron network), trí tuệ nhân tạo (artificial intelligence) chẳng hạn để phân loại có phải là vùng biển số hay không.

2.2.1. Phương pháp chuyển đổi Hough

Dò đặc trưng biên ngang, dọc: làm nổi bật các viền bao của tất cả các đối tượng trong ảnh trong đó có viền bao biển số. Phương pháp sử dụng các bộ lọc gradient để trích được các đặc trưng cạnh biên này. Nghiên cứu này sử dụng bộ lọc Sobel để tiến hành dò. Dùng chuyển đổi Hough tìm các đoạn thẳng ngang dọc trên cơ sở của ảnh nhị phân cạnh thu được từ bước trên. Tách các đoạn thẳng ngang, dọc có thể là cạnh của biển số.

Trích ứng viên biển số: thành lập các hình chữ nhật là ứng viên cho biển số với tiêu chí cụ thể là các bộ 4 đoạn thẳng thu được sẽ qua đánh giá về kích thước, tỉ lệ chiều rộng trên chiều cao so với một ngưỡng nào đó.

Ưu điểm: độ chính xác cao, không phụ thuộc vào màu sắc của biển số xe.

Nhược điểm: Độ phức tạp tính toán khá cao. Khi ảnh có thêm nhiều đối tượng khác thì khối lượng tính toán tăng lên rất nhiều do mục đích là phải xác định được vùng con nào chứa biên số xe và phụ thuộc rất lớn vào bước trích đặc trưng biên cạnh dẫn đến là các đoạn thẳng ứng viên thu được ngắn thường ngắn hơn nhiều so với chiều dọc cũng như chiều ngang của biên số.

2.2.2. Phương pháp hình thái học

Nhóm tác giả Chirag N. Paunwala, 2012 đại diện cho phương pháp này, với kết quả nhận dạng rất tốt 99.5%.

Nội dung của phương pháp: Dựa vào đặc trưng quan trọng là biên số xe máy có độ sáng (tức mức xám khi chuyển bức ảnh về dạng xám) là tương đối khác so với các vùng khác trong bức ảnh, cũng như sự phân bố mức xám là khá đồng đều trên biên số và vì vậy khi được nhị phân hoá, vùng biên số là một đối tượng có đặc thù hình thái, có thể phân biệt được với các vùng khác. Như vậy các bước thực hiện là:

Xác định ngưỡng xám. Thực chất là không có phương pháp nào chọn cho đúng ngưỡng xám để thực hiện. Thay vào đó, ngưỡng xám sẽ được quét trong một khoảng nào đó. Thông qua lược đồ xám ta nhận thấy vùng biên số thường sẽ có độ sáng tương đối lớn (từ 130-200) vì vậy ta sẽ xác định ngưỡng xám cần chọn sẽ thuộc vùng này nhờ đó ta sẽ giảm được thời gian lặp tìm ngưỡng xám.

- Nhị phân hoá ảnh xám đầu vào với ngưỡng xám đã xác định.
- Lọc các nhiễu gây ảnh hưởng xấu tới đối tượng biên số.
- Gắn nhãn cho các đối tượng trong ảnh nhị phân thu được.
- Trích ra các đối tượng ứng viên biên số theo tiêu chí cụ thể của biên số xe về chiều cao, chiều rộng, tỉ lệ các cạnh, diện tích, trọng tâm, số điểm cắt...



Hình 2-1. Ảnh xám và lược đồ xám của ảnh

Việc nghiên cứu ra các thuật toán rất quan trọng trong ngành khoa học máy tính nói chung, công nghệ thông tin và lập trình phần mềm nói riêng. Các thuật toán được trình bày trong phần này sẽ giúp chọn ra một địa điểm có khoảng cách người dùng ngắn nhất trong tập hợp các địa điểm trong cơ sở dữ liệu một cách nhanh và chính xác nhất khi người dùng liên tục thay đổi vị trí của mình giúp cho ứng dụng có thể vận hành một cách tự động. Thuật toán được nghiên cứu giúp giải quyết các vấn đề được đưa ra trong bài toán sau đây.

2.3. Phương pháp nhận dạng ký tự trong biển số xe

Phương pháp phổ biến nhất để nhận dạng ký tự là sử dụng mạng nơ-ron (hoặc SVM, K-NN,...), tức là huấn luyện cho máy tính để nhận dạng các ký tự. Tuy nhiên do số lượng ký tự trên biển số là không nhiều nên để đảm bảo tốc độ xử lý, chúng ta cũng có thể sử dụng phương pháp Hình thái học để giải quyết khâu này bởi vì các ký tự đều có những đặc điểm hình thái đặc biệt có thể phân biệt với nhau chẳng hạn như “0” có lỗ trống ở giữa, “8” có 2 lỗ trống hay “X” đối xứng 2 trục dọc và ngang... Khâu này được thực hiện trên cơ sở xây dựng cây nhị phân tối ưu của các đặc điểm hình thái nên đảm bảo tính khoa học và tính chính xác cao. Thuật toán cơ bản của bước này như sau:

- Quan sát chọn ra các đặc tính phân biệt ký tự để xây dựng ma trận đặc tính.
- Xây dựng cây nhị phân tối ưu từ ma trận đặc tính và ttpaj ký tự thu được.
- Quan sát cây nhị phân, kiểm tra số đặc tính như vậy đã đủ để nhận dạng chưa, thiếu (dư) thì phải bổ sung (bỏ đi) và quay lại bước đầu tiên.
- Tiến hành nhận dạng các ký tự trên cơ sở cây nhị phân tối ưu tìm được.

2.4. Phạm vi nghiên cứu và hướng giải quyết

Trong đồ án này, công việc cần phải giải quyết vấn đề phát hiện vùng chứa biển số xe và nhận dạng ký tự trong biển số. Với bài toán phát hiện vùng chứa biển số cách tiếp cận theo phương pháp hình thái học để phát hiện vùng biển số cho các biển đăng ký xe của Việt Nam. Bài toán nhận dạng ký tự sẽ sử dụng mô hình mạng Noron tích chập để tiến hành nhận dạng.

Một số đặc điểm để nhận dạng biển số xe tại Việt Nam.

Tiêu chuẩn về kích thước (theo quy định của Bộ Giao Thông Vận Tải)

Ở mỗi nước thường có tiêu chuẩn về kích thước nhất định. Đối với nước ta biển số xe qui định khá đồng đều cho mỗi loại xe, tỷ lệ chiều dài, rộng cho mỗi loại xe là như nhau. Quy định về kích thước như sau:

- Biển ô tô:
 - Chiều cao: 110 mm. Chiều rộng: 470 mm (biển dài).
 - Chiều cao: 200 mm. Chiều rộng: 280 mm (biển vuông).
- Biển xe máy:
 - Chiều cao: 140 mm. Chiều rộng: 190 mm.

Như vậy, tỉ lệ Chiều cao / Chiều rộng là:

- $0.18 < \text{Chiều cao/Chiều rộng} < 0.3$ (biển số có 1 hàng).
- $0.6 < \text{Chiều cao/Chiều rộng} < 0.85$ (biển số xe có 2 hàng).

Từ các đặc điểm này, ta có thể xác định được vùng nào có khả năng là biển số theo ràng buộc về kích thước.

Tiêu chuẩn về ký tự

Theo đo đạc trên biển số thực tế, mỗi ký tự thường có tỷ lệ kích thước về chiều rộng, chiều cao tương ứng với chiều dài và rộng của biển số xe. Ví dụ, chiều cao của mỗi ký tự luôn nhỏ hơn 85% chiều cao của biển số xe và luôn lớn hơn 60% chiều cao của biển xe đối với biển số xe có một hàng, với biển số xe có hai hàng thì chiều cao mỗi ký tự không quá 50% chiều cao của biển số xe. Chiều rộng của ký tự thường không lớn hơn 20% chiều cao của mỗi ký tự. Mỗi ký tự của biển số xe được xem như là một vùng liên thông con hay contour con (bao đóng). Do đó, chúng ta có thể đếm số contours con thỏa mãn tính chất đó là ký tự. Ở nước ta chỉ có số ký tự trên mỗi biển số xe nằm trong khoảng 6 đến 9 ký tự.

Từ những phân tích trên, có thể tóm tắt phương pháp thực hiện của chương trình nhận dạng như sau:

Bước 1: Ảnh đầu vào là ảnh màu BGR, tiền xử lý bằng các thuật toán xử lý ảnh.

Bước 2: Tìm các contour trên ảnh xe. (Mỗi contour là 1 vùng bao kín, do vùng biển số là 1 vùng bao kín nên sẽ tương ứng với một contour).

Bước 3: Lọc các contour theo các tiêu chí như kích thước, góc, tỉ lệ, số ký tự...

Bước 4: Xử lý kết quả đầu ra để lấy vùng biển số.

Bước 5: Tách ký tự trên vùng biển số tìm được.

Bước 6: Đưa tập ký tự đã tách vào mạng Noron để nhận dạng.

2.5. Kết luận chương

Dựa trên các nội dung thu thập và phân loại về hệ thống nhận dạng biển số xe giúp chúng ta có định hướng rõ ràng và đưa ra các phương pháp áp dụng nhận dạng.

Đã giải quyết được vấn đề phát hiện vùng chứa biển số xe và nhận dạng ký tự trong biển số. Với bài toán phát hiện vùng chứa biển số cách tiếp cận theo phương pháp hình thái học để phát hiện vùng biển số cho các biển đăng ký xe. Bài toán nhận dạng ký tự sẽ sử dụng mô hình mạng nơron tích chập để tiến hành nhận dạng.

CHƯƠNG 3. ÁP DỤNG MẠNG NƠON TÍCH CHẬP TRONG NHẬN DẠNG KÝ TỰ

3.1. Nhận dạng ký tự trong bài toán nhận dạng biển số xe

Nhận dạng ký tự là một bước quan trọng và chủ chốt trong bài toán nhận dạng biển số xe. Vấn đề nhận dạng ký tự nói chung và nhận dạng ký tự viết tay nói riêng là một thách thức lớn đối với các nhà nghiên cứu, vì sự phức tạp của việc nhận dạng ký tự viết tay phụ thuộc nhiều vào phong cách viết và cách thể hiện ngôn ngữ của người viết, còn nhận dạng ký tự trên biển số xe phụ thuộc nhiều vào các yếu tố như chất lượng ảnh, thời tiết, nhiệt độ...

Nhận dạng ký tự đã được nghiên cứu và ứng dụng thành công trong nhiều lĩnh vực như: Nhiều bưu điện đã áp dụng hệ thống phân loại thư tự động dựa trên máy đọc bì thư. Các ngân hàng đọc nội dung của séc để chống rửa tiền, gian lận và cả phát hiện khủng bố. Google đã phát triển ứng dụng google dịch, thay vì việc nhập ký tự trên bàn phím chúng có thể nhận dạng ký tự qua camera điện thoại và dịch. Với giải thuật tối ưu, tốc độ xử lý nhanh và mức độ chính xác cao cũng như đã ứng dụng thành công trong nhiều đó là lý do chúng ta áp dụng phương pháp nhận dạng ký tự vào bài toán nhận dạng biển số xe trong luận văn này.

3.2. Giai đoạn xử lý ảnh

3.2.1. Cách phân loại ảnh

Ảnh nhị phân: Giá trị xám của tất cả các điểm ảnh chỉ nhận giá trị 1 hoặc 0 như vậy mỗi điểm ảnh trong ảnh nhị phân được biểu diễn bởi 1 bit. Ảnh xám: Giá trị xám nằm trong $[0, 255]$ như vậy mỗi điểm ảnh trong ảnh nhị phân được biểu diễn bởi 1byte.

Ảnh màu:

Hệ màu RGB: Một pixel được biểu diễn bằng 3 giá trị (R, G, B) trong đó R, G, B là một giá trị xám và được biểu biểu diễn bằng 1 byte. Khi đó ta có một ảnh 24 bits.

$$P(x, y) = (R, G, B)$$

Hệ màu CMY: là phần bù của hệ màu RGB, thường được dùng trong máy in.

$$(C, M, Y) = (1, 1, 1) - (R, G, B)$$

$$\text{Hay } C+R=M+G=Y+B=1$$

Hệ màu CMYK: trong đó K là độ đậm nhạt của màu.

$$K = \min(C, M, Y)$$

$$P(x, y) = (C-K, M-K, V-K, K).$$

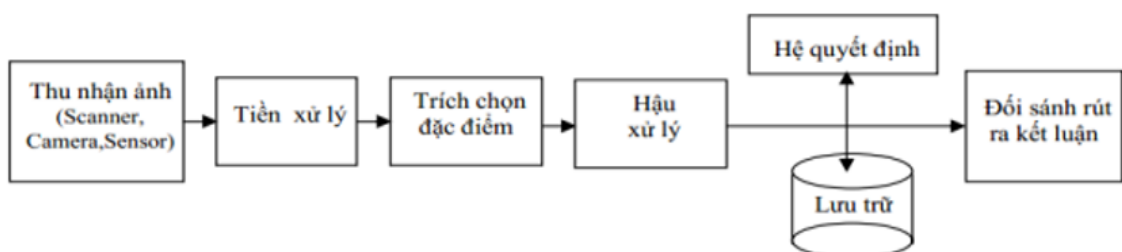


Hình 3-1. Ảnh màu, ảnh mức xám và ảnh nhị phân

3.2.2. Các giai đoạn xử lý ảnh

Giai đoạn xử lý ảnh trong nhận dạng biến số xe gồm các bước sau:

- **Thu nhận ảnh:** ảnh được thu từ nhiều nguồn khác nhau: máy ảnh, máy quay phim, máy quét, ảnh vệ tinh biến đổi thông tin hình ảnh về các cấu trúc được lưu trữ trong máy tính, có thể hiển thị ra các thiết bị ngoại vi.
- **Tiền xử lý** (Image Processing): quá trình sử dụng các kỹ thuật xử lý ảnh để làm ảnh tốt lên theo mục đích sử dụng.
- **Phân đoạn và biểu diễn ảnh.**
- **Nhận dạng và nội suy ảnh.**



Hình 3-2. Các bước xử lý ảnh cơ bản

3.2.3. Xác định vùng chứa biển số

a. Phân ngưỡng vùng đối tượng

Ý tưởng: nhận thấy các vùng biển số là các vùng đặc trưng có hình chữ nhật, có tỉ lệ kích thước theo tiêu chuẩn. Vì vậy sau khi tách ngưỡng và tìm vùng (EmguCV hỗ trợ findContours) ta tìm những vùng có tỷ lệ theo kích thước giống với tỉ lệ của biển số xe. Để có thể lọc các vùng biển số một lần nữa thì ta tiếp tục tìm số vùng con trên vùng biển số và so sánh với số lượng ký tự của biển số.

Gồm 2 bước:

- Bước 1: Tìm vùng biển số với ảnh xám đầu vào thực hiện phân ngưỡng (hoặc dò biên canny) và tìm đối tượng. Tách các đối tượng có tỉ lệ rộng/dài trong phạm vi biển số để làm nguồn cho bước 2.
- Bước 2: Tách riêng từng ký tự trên biển số: Tiến hành tìm đối tượng trên nguồn do bước 1 cung cấp sau đó so sánh số đối tượng nhận được có trùng với số ký tự tên các biển số xe không. Nếu khớp với tiêu chuẩn thì đó là những vùng biển số có khả năng là biển số xe cao nhất. Tách riêng từng ký tự để nhận dạng.

b. Một số đặc điểm xe ở Việt nam

Ở mỗi nước thường có tiêu chuẩn về kích thước nhất định. Đối với nước ta, biển số xe qui định khá đồng đều cho mỗi loại xe, tỷ lệ chiều dài, rộng cho mỗi loại xe là như nhau:

Đối với loại xe có một hàng ký tự thì tỉ lệ dài/rộng là: $3.5 \leq W / H \leq 4.5$.

Đối với loại xe có hai hàng ký tự thì tỉ lệ dài/rộng là: $0.8 \leq W / H \leq 1.4$.

Từ những đặc điểm này, ta có thể xác định được các vùng con thỏa mãn thì khả năng chứa biển số là rất cao.

Mỗi ký tự thường có tỷ lệ kích thước chiều rộng, chiều cao tương ứng với chiều dài và rộng của biển số xe. Ví dụ, chiều cao của mỗi ký tự luôn nhỏ hơn 85% chiều cao của biển số xe và luôn lớn hơn 33% chiều cao của biển xe. Còn chiều

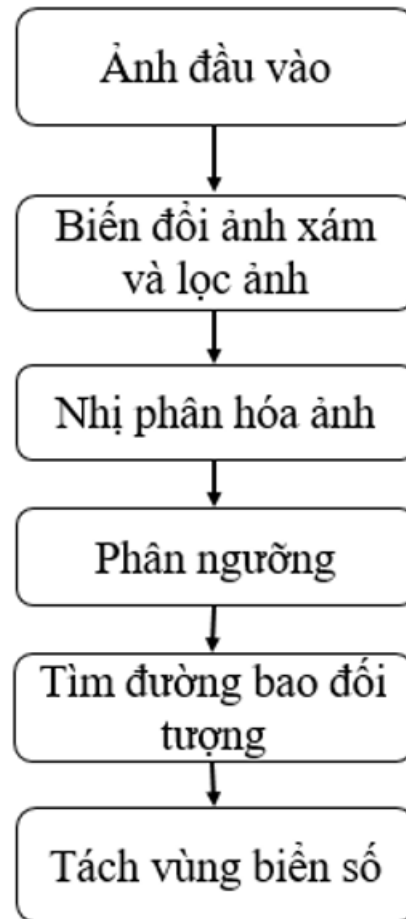
rộng của ký tự không lớn hơn 20% chiều dài của biển số xe. Mỗi ký tự của biển số xe được xem như là một vùng liên thông con. Do đó, chúng ta có thể đếm vùng liên thông con thỏa mãn tích chất đó là ký tự. Chú ý số ký tự trên biển số xe là từ 6 đến 10 ký tự. ở nước ta chỉ có số ký tự trên mỗi biển số xe nằm trong khoảng 6 đến 9 ký tự.



Hình 3-3. Một số loại biển số xe thông dụng

c. Phát hiện vùng chưa biển số

Sơ đồ các bước phát hiện vùng chứa biển số xe.



Hình 3-4. Các bước phát hiện vùng chứa biển số xe

Bước 1: Biến đổi ảnh xám và lọc ảnh

Ảnh đầu vào là một ảnh bất kỳ, được chuyển về ảnh có 256 mức xám và tiến hành lọc ảnh giảm nhiễu.

Bước 2: Nhị phân hóa ảnh

Việc sử dụng ảnh mức xám không làm giảm đi tính đa năng của ứng dụng. Trên thực tế, ảnh mức xám vẫn được sử dụng nhiều, và nhiều thiết bị ghi hình cũng có khả năng tự chuyển ảnh màu thành ảnh mức xám. Tuy nhiên, nếu để ảnh mức xám thì việc phát hiện biên không hiệu quả, vì sự thay đổi liên tục của các mức xám làm cho việc xác định biên không phải dễ dàng, và việc tìm ra các vùng liên tục của biên khá hạn chế. Vì vậy, chúng ta thực hiện chuyển ảnh về dạng nhị phân để thực hiện việc lấy biên nhanh hơn.

Bước 3: Tiến hành phân ngưỡng hoặc phát hiện biên.

Có hai quá trình phân ngưỡng đó là phân ngưỡng tự động và phân ngưỡng không tự động.

Khi ta tiến hành phân ngưỡng sẽ làm hiện rõ vùng biên số, trong giai đoạn này ta có thể sử dụng các phương pháp phát hiện biên.

Bước 4: Tìm đường bao đối tượng.

Khi có ảnh thu được ở bước 3 chúng ta tiến hành tìm các vùng biên là đối tượng riêng để trích các vùng và lấy thông số của các vùng như là tọa độ điểm của các cạnh, diện tích của vùng.

Bước 5: Tách vùng biên số.

Tìm đường bao đối tượng với tiêu chí tỉ lệ $3.5 \leq W/H \leq 4.5$ hoặc $0.8 \leq W/H \leq 1.4$ để tìm ra những vùng có thể là biên số nhất.

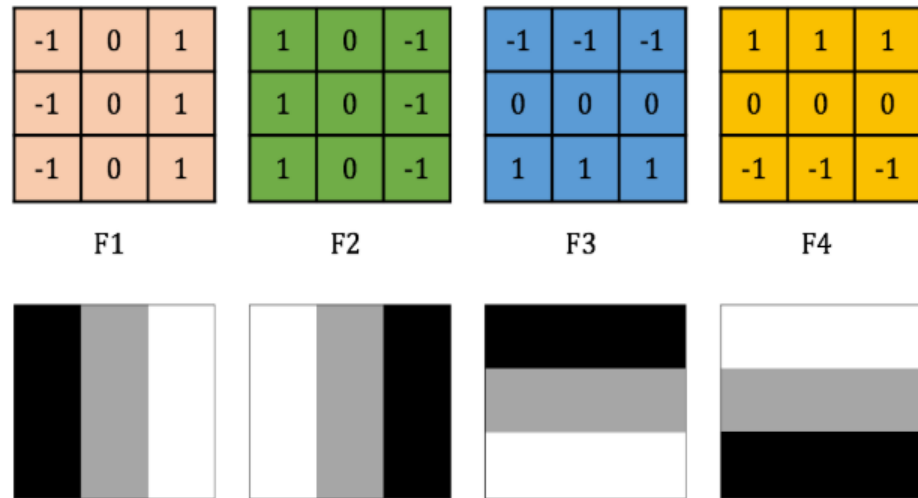
Tiến hành cắt các vùng có thể là biên số nhất trên ảnh để làm nguồn cho giai đoạn cắt ký tự để làm mẫu cho các quá trình nhận dạng ký tự.

3.3. Xây dựng mô hình nhận dạng ký tự

Sau khi đã tách được các ký tự dưới dạng ảnh từ biên số xe, bước tiếp theo chúng ta cần nhận dạng các ký tự từ ảnh chuyển về text. Để nhận dạng được các ký tự có rất nhiều phương pháp nhận dạng, có thể là KNN, SVM, mạng nơron lan truyền ngược,... Ở đây chúng ta sử dụng mạng nơron lan tích chập trong nhận dạng ký tự số nguyên do mạng nơron tích chập có độ chính xác cao, và hiệu năng tốc độ xử lý tính toán nhanh hơn các mạng trước đó.

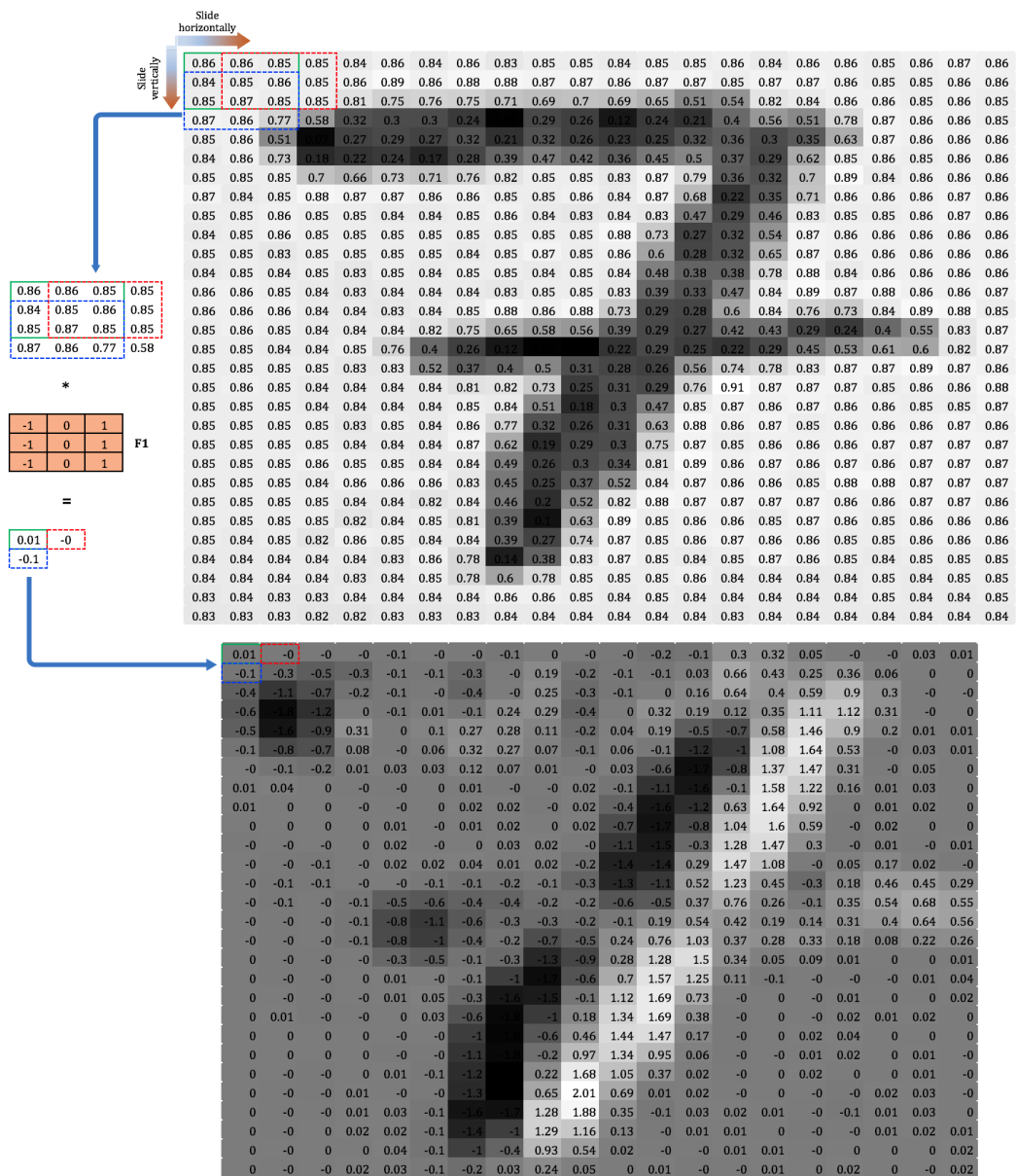
CNN được sử dụng để phân loại tập các ảnh viết tay của các số từ 00 tới 99. Đầu vào là những bức ảnh trắng đen (Gray Scale) và được biểu diễn bởi một ma trận các điểm ảnh với kích thước cố định $h \times w \times w$. Lớp tích chập đầu tiên của CNN sử dụng 44 bộ lọc kích thước $3 \times 3 \times 3$: F1F1, F2F2, F3F3, F4F4 với giá trị tương

ứng như trong hình 1. Các giá trị tại mỗi ô của các bộ lọc có thể được biểu diễn bởi màu sắc tương ứng với Đen (-1), Xám (0), Trắng (1) như trong hình dưới đây.



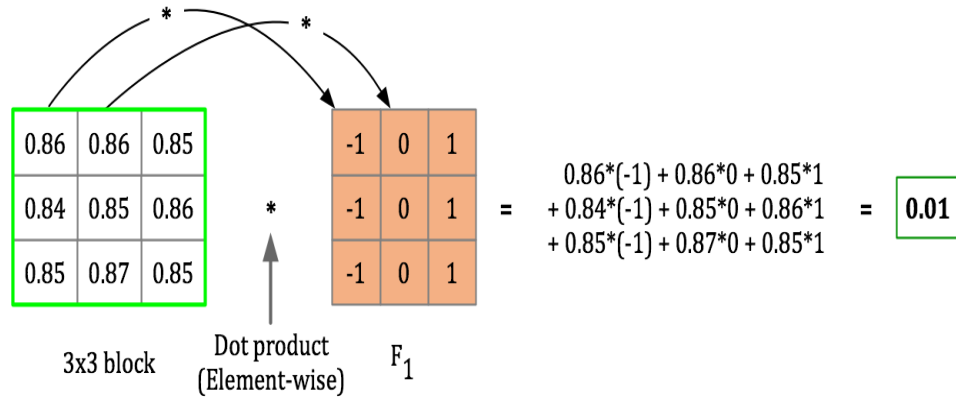
Hình 3-5. Bộ lọc được sử dụng trong lớp tích chập đầu tiên là các ma trận kích thước 3×3 của $-1, 0$ và 1

Để minh họa cho phép nhân chập, chúng ta sử dụng đầu vào là một bức ảnh viết tay của số 77, biểu diễn dưới dạng ma trận $30 \times 22 \times 30 \times 22$ và áp dụng riêng biệt từng bộ lọc ở trên. Phép nhân tích chập được thực hiện bằng cách trượt ma trận lọc $3 \times 3 \times 3$ trên ma trận ảnh đầu vào $32 \times 22 \times 32 \times 22$ (bộ lọc dịch sang phải/ xuống dưới 11 cột/ hàng mỗi một lần trượt) cho đến khi nó đi qua hết tất cả các vùng kích thước $3 \times 3 \times 3$. Việc trượt ma trận lọc trên ma trận đầu vào được gọi là “chập” (convolving). Như minh họa trong hình 3.2, ma trận F1F1 được chập với từng vùng (block - region) điểm ảnh kích thước $3 \times 3 \times 3$ của ảnh đầu vào. Tại mỗi vị trí di chuyển của ma trận F1F1, giá trị đầu ra được tính bằng tích chập (dot-product) của ma trận F1F1 với vùng bao phủ tương ứng.



Hình 3-6. Nhân chập bộ lọc F_1 với ma trận ảnh đầu vào của số 7

Ô đầu tiên (0,0)(0,0) của ma trận đầu ra (giá trị 0.010.01) ra là kết quả của phép nhân chập giữa ma trận $F_1 F_1$ với góc trái trên cùng của ma trận đầu vào và được tính như sau:



Từ các ma trận đầu ra kích thước $28 \times 2028 \times 20$, chúng ta thấy được cả bốn bộ lọc F1F1, F2F2, F3F3 và F4F4 đều được sử dụng để phát hiện cạnh trong bức ảnh (thể hiện bởi những điểm ảnh sáng hơn)

- F1: Phát hiện cạnh đứng phải.
- F2: Phát hiện cạnh đứng trái.
- F3: Phát hiện cạnh ngang dưới.
- F4: Phát hiện cạnh ngang trên.



Hình 3-7. Ví dụ về bộ lọc cạnh (đứng phải, đứng trái, ngang dưới, ngang trên) với đầu vào là ảnh số viết tay.

3.4. Thuật toán

3.4.1. Tách dòng

Bước 1: Xác định giới hạn dưới của dòng.

Bước 2: Bắt đầu duyệt từ giới hạn trên (đỉnh) vừa tìm thấy của dòng (0, top_line).

Bước 3: Tương tự như xác định giới hạn trên, chúng ta duyệt hết chiều rộng của ảnh trên cùng một giá trị y. Nếu duyệt hết dòng mà không tìm thấy ký tự pixel đen nào thì ghi nhận y-1 là giới hạn dưới của dòng (bottom_line). Dừng duyệt. Tăng số dòng lên (lines++).

Nếu chưa tìm thấy bottom_line, tiếp tục duyệt đến dòng tiếp theo (tăng y, reset x=0).

Bắt đầu từ giới hạn dưới y (bottom_line) vừa tìm thấy sau cùng, lặp lại các bước a,b để xác định các giới hạn của các dòng tiếp theo, cho đến khi duyệt hết chiều cao của ảnh thì dừng, quá trình xác định dòng ký tự hoàn tất.

3.4.2. Tách ký tự

Bước 1: Bắt đầu từ ký tự đầu tiên của hàng trên cùng với giá trị x đầu tiên.

Bước 2: Quét hết chiều rộng với một giá trị y.

Nếu phát hiện điểm đen đánh dấu y như là đỉnh của hàng đầu tiên

Nếu không xét điểm tiếp theo

Bước 3: Bắt đầu từ giới hạn trên của ký tự phát hiện được và giá trị x đầu tiên.
(0, giới hạn trên ký tự)

Bước 4: Quét đến giới hạn dưới của dòng, giữ nguyên x.

Nếu phát hiện điểm đen đánh dấu x là phía trái của ký tự.

Nếu không xét điểm tiếp theo.

Nếu không thấy điểm đen nào tăng x và khởi động lại y để xét đường thẳng đứng tiếp theo.

Bước 5: Bắt đầu từ phía trái của ký tự tìm thấy

3.4.3. Ánh xạ vào ma trận

Bước 1: Đối với chiều rộng:

Khởi tạo với 19 phần tử tương ứng

Ánh xạ điểm đầu $(0,y)$ và điểm cuối (C_rong,y) của ảnh kí tự tương ứng với giá trị đầu $(0,y)$ và giá trị cuối $(10,y)$ của ma trận.

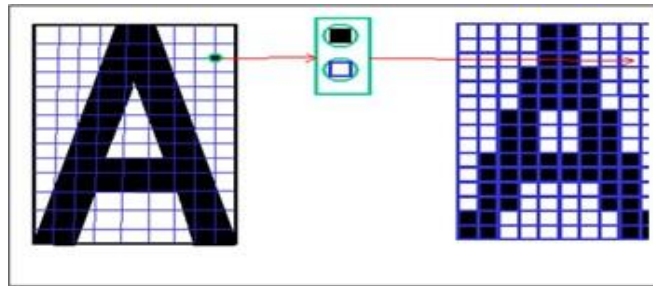
Chia nhỏ chiều rộng thành 19 giá trị tương ứng

Bước 2: Đối với chiều cao:

Khởi tạo với 29 phần tử tương ứng

Ánh xạ điểm đầu $(x,0)$ và điểm cuối (x,C_cao) của ảnh kí tự tương ứng với giá trị đầu $(x,0)$ và giá trị cuối $(x,29)$ của ma trận

Chia nhỏ chiều cao thành 19 giá trị tương ứng



Hình 3-8. Quá trình chia lưới kí tự

Để đưa giá trị vào mạng noron, chúng ta cần chuyển ma trận điểm ảnh sang ma trận giá trị.

Thuật toán:

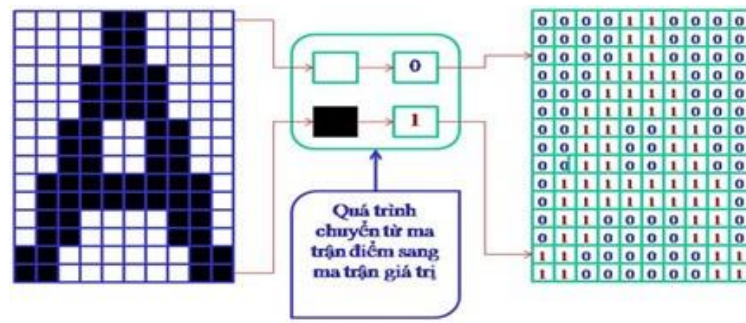
Bước 1: Bắt đầu từ phần tử $(0,0)$

Bước 2: Tăng x giữ nguyên giá trị y cho tới khi bằng chiều rộng của ma trận

Ánh xạ mỗi phần tử tới một phần tử của mảng tuyến tính

- Nếu là điểm đen thì nhận giá trị bằng 1
- Ngược lại nhận giá trị bằng 0
- Nếu $x = \text{chiều rộng}$ thì khởi động lại x và tăng y

Lặp lại cho tới khi $(x,y) = (C_Rong, C_Cao)$



Hình 3-9. Quá trình ánh xạ từ ma trận điểm sang ma trận giá trị

3.4.4. Huấn luyện mạng nơron

Bước 1: Xây dựng mạng tương ứng với mô hình tham số

Bước 2: Khởi tạo giá trị trọng số với giá trị ngẫu nhiên. Nạp file huấn luyện (cả ảnh đầu vào và ảnh đầu ra mong muốn)

Bước 3: Phân tích ảnh và ánh xạ tất cả kí tự tìm thấy vào các mảng một chiều

Bước 4: Đọc giá trị đầu ra mong muốn từ file và lưu trữ riêng biệt

Bước 5: Với mỗi kí tự:

- Tính toán giá trị đầu ra của mạng Feed ForWard
- So sánh với giá trị đầu ra mong muốn tương ứng với từng kí tự và tính toán lỗi
- Truyền ngược giá trị từ đầu và với mỗi liên kết điều chỉnh trọng số liên kết

Bước 6: Chuyển sang kí tự tiếp theo và lặp lại “6” cho tới khi hết các kí tự

Bước 7: Tính toán trung bình lỗi cho tất cả các kí tự

Bước 8: Lặp lại từ bước 6 đến 8 cho tới khi đạt số đưa vào của số lần lặp tối đa

Với phạm vi lỗi đạt đến ngưỡng. Nếu như vậy thì bỏ lặp lại, ngược lại tiếp tục lặp lại

3.4.5. Nhận dạng ký tự

Bước 1: Nạp file ảnh

Bước 2: Phân tích ảnh cho các dòng kí tự

Bước 3: Với mỗi dòng tách các kí tự liên tiếp

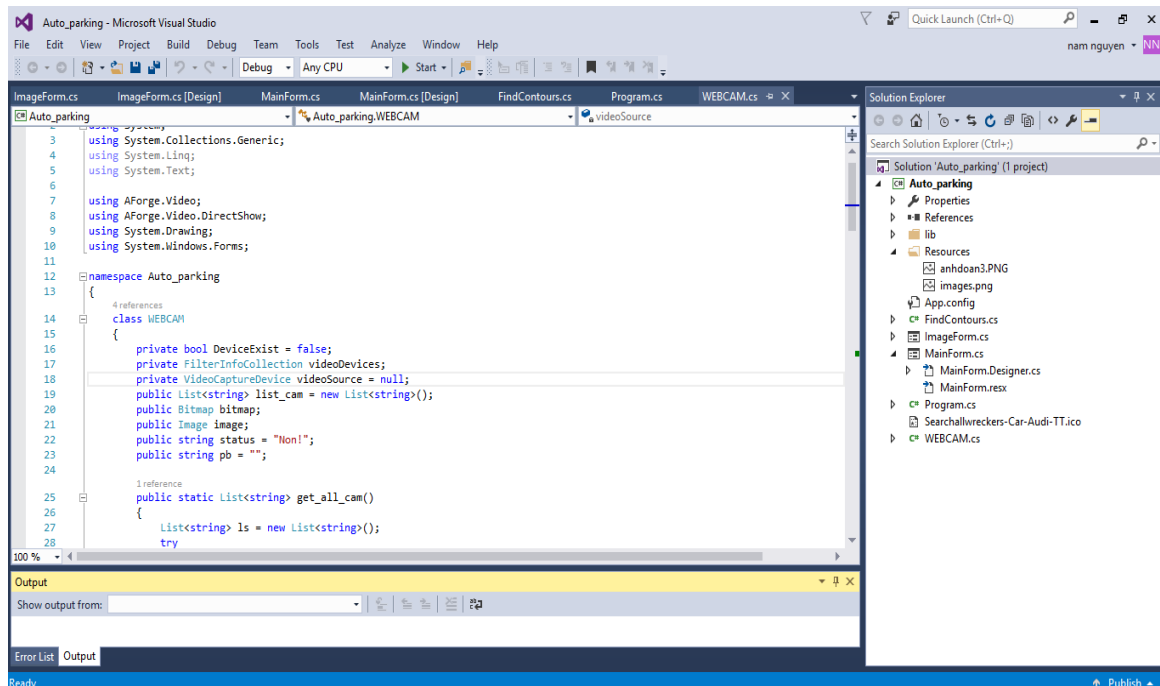
- Phân tích và xử lý ảnh kí tự cho việc ánh xạ vào một vectơ đầu vào
- Đưa giá trị vector đầu vào cho mạng nơron và tính toán giá trị đầu ra

3.5. Cài đặt ứng dụng

3.5.1. Môi trường cài đặt và các yêu cầu phần cứng phần mềm

Phần mềm Visual studio 2015: Visual Studio là một môi trường tích hợp từ Microsoft. Nó được sử dụng để phát triển chương trình máy tính cho Microsoft Windows, cũng như các trang web, các ứng dụng web và các dịch vụ web. Visual Studio hỗ trợ nhiều ngôn ngữ lập trình khác nhau và cho phép các lập trình viên biên tập mã và gỡ lỗi để hỗ trợ (mức độ khác nhau) với hầu hết các ngôn ngữ lập trình, cung cấp một dịch vụ ngôn ngữ cụ thể tồn tại. . Ứng dụng này được xây dựng trong ngôn ngữ bao gồm C, C ++, VB, C#...

Tính năng chính của chương trình Visual Studio là biên tập code, sửa lỗi và thiết kế Window Forms...



Hình 3-10. Giao diện Visual studio 2015

Thư viện xử lý ảnh EmguCV: OpenCV (Open Computer Vision) là một thư viện mã nguồn mở chuyên dùng để xử lý các vấn đề liên quan đến thị giác máy tính. Nhờ một hệ thống các giải thuật chuyên biệt, tối ưu cho việc xử lý thị giác máy tính, vì vậy tính ứng dụng của OpenCV là rất lớn. EmguCV là một cross platform.NET, một thư viện xử lý hình ảnh mạnh dành riêng cho ngôn ngữ C#, cho phép gọi được chức năng của OpenCV từ .NET.

Lợi thế của EmguCV là được viết hoàn toàn bằng C# ,có thể chạy trên bất kỳ nền tảng hỗ trợ nào.

Tính ứng dụng của EmguCV là rất lớn, có thể kể đến như nhận dạng ảnh (nhận dạng khuôn mặt, các vật thể ...),xử lý ảnh (khử nhiễu, điều chỉnh độ sáng ...), nhận dạng cử chỉ và còn rất nhiều ứng dụng khác nữa.

3.5.2. Giao diện chương trình chính



Hình 3-11. Giao diện chương trình chính

3.6. Kết quả thực nghiệm nhận dạng biển số xe

Chạy ứng dụng nhận dạng biển số, đưa lần lượt 376 ảnh dữ liệu về xe để kiểm tra kết quả phát hiện biển số, tách ký tự và nhận dạng ký tự của chương trình, ghi nhận kết quả.

Tỉ lệ biển nhận dạng đúng vùng biển số: 372/376 ~ 98%



Biển bị bóng, phản chiếu

Biển số bị tối, thiếu ánh sáng

Biển dơ, mù, ố vàng

Hình 3-12. Một số biển không phát hiện được biển số

Tỉ lệ biển tách đúng và đầy đủ ký tự: 330/335 ~ 98%



Hình 3-13. Một số biển không tách đúng ký tự, chữ số bị dính với các vật bên ngoài như đinh ốc, ký tự bị mờ nét, mất nét, loang lổ,...



Hình 3-14. Một số biển số xe nhận dạng sai hặc không thể nhận dạng



Hình 3-15. Mẫu biển số xe nhận dạng chuẩn

Kết quả nhận dạng chung của ứng dụng từ khâu phát hiện biển số, đến tách ký tự và nhận dạng ký tự đạt xấp xỉ 65% với dữ liệu có nhiều ảnh không đạt tiêu chuẩn như bị bóng mờ, quá dơ, nhòe,... Với ảnh biển số sạch sẽ, rõ ràng, không chứa các phụ kiện gắn trên biển, tỉ lệ nhận dạng có thể đạt tới hơn 70%.

3.7. Kết luận chương

Demo nhận dạng với phương pháp trình bày trong đồ án này đã thu được một số điểm sau:

1. Huấn luyện mạng nơron có thành công hay không phụ thuộc nhiều vào các giá trị khởi tạo ban đầu. Nếu lựa chọn được giá trị tối ưu thì việc huấn luyện sẽ nhanh đạt được kết quả.
2. Chất lượng của quá trình huấn luyện phụ thuộc nhiều vào chất lượng của tập mẫu. Nếu tập mẫu càng nhiều, đầy đủ, thì chất lượng nhận về tri thức của mạng càng cao, tuy nhiên thời gian huấn luyện sẽ lâu hơn.
3. Thời gian huấn luyện phụ thuộc nhiều vào các lựa chọn tham số đầu vào tại các nút ẩn, nút ra, nút vào.

Ưu điểm: thuật toán cài đặt nhanh, tìm vùng biên số và cách ly ký tự với tỉ lệ thành công cao (ở những biên số thông thường), tìm được ở những ảnh tự nhiên, vùng biên số bị nghiêng. Bên cạnh đó, còn tồn tại một số điểm hạn chế chưa giải quyết được như:

- Tỉ lệ nhận dạng còn phụ thuộc nhiều vào điều kiện ánh sáng, phản chiếu hoặc che bóng.
- Với những biên số có đường viền phức tạp thì mức độ nhận dạng không cao. Chỉ giới hạn ảnh chụp trong góc 40 độ để nhận dạng được tốt, nếu góc lớn hơn thì khả năng nhận dạng được giảm.
- Ảnh xe trong bóng tối hay ban đêm không có đèn flash thì khả năng nhận dạng thấp do thiếu sáng.

Bên cạnh đó, còn tồn tại một số điểm hạn chế chưa giải quyết được như:

- Tỉ lệ nhận dạng còn phụ thuộc nhiều vào điều kiện ánh sáng, phản chiếu hoặc che bóng.
- Với những biên số có đường viền phức tạp thì mức độ nhận dạng không cao. Chỉ giới hạn ảnh chụp trong góc 40 độ để nhận dạng được tốt, nếu góc lớn hơn thì khả năng nhận dạng được giảm.
- Ảnh xe trong bóng tối hay ban đêm không có đèn flash thì khả năng nhận dạng thấp do thiếu sáng.

KẾT LUẬN

Luận văn “*Nghiên cứu về mạng nơron tích chập và ứng dụng cho bài toán nhận dạng biển số xe*” với mục đích tìm hiểu bài toán giám sát, quản lý các phương tiện giao thông một cách tự động đã đạt được những kết quả chính sau:

Lý thuyết:

Về mặt lý thuyết, luận văn đã giới thiệu về mạng nơron tích chập cũng như một số phương pháp nhận dạng ảnh, ký tự; luận văn cũng đưa ra các nghiên cứu một số thuật toán tính áp dụng cho bài toán được đưa ra và áp dụng các thuật toán đó vào ứng dụng.

Ứng dụng:

Trên cơ sở nghiên cứu lý thuyết đã test thành công các dữ liệu đầu vào bằng ứng dụng xây dựng trên nền tảng .NET cho ra các kết quả khả quan. Hiện tại ứng dụng đang trong giai đoạn thử nghiệm và phát triển thêm modul để tiếp thu những ý kiến đóng góp của người dùng giúp cho ứng dụng được hoàn thiện hơn.

Hướng nghiên cứu tiếp theo:

Nâng cao hiệu quả chương trình, tách ly các ký tự trong biển số trong các trường hợp biển số bị nhiễu nhiễu, mất mát thông tin do nhiễu từ điều kiện môi trường, tìm vùng biển số trong ảnh có độ tương phản giữa biển số và nền thấp. Đặc biệt là biển xe có nền màu đỏ chữ trắng.

Phát triển chương trình thành module phần cứng. Có khả năng tương thích với các thiết bị quan sát như camera.

Nghiên cứu theo hướng một ứng dụng cụ thể như : giám sát phương tiện giao thông, xử lý vi phạm giao thông, quản lý xe tại các bãi giữ xe, các kho vật tư...

DANH MỤC TÀI LIỆU THAM KHẢO

Tài liệu Tiếng Việt

- [1] Trần Thúy Anh (chủ biên) (2011), Du lịch văn hóa những vấn đề lý luận và nghiệp vụ, Nxb Giáo dục Việt Nam, tr.37.

Tài liệu Tiếng Anh:

- [2] Y. LeCun and Y. Bengio. "Convolutional networks for images, speech, and time-series." In M. A. Arbib, editor, *The Handbook of Brain Theory and Noron Networks*. MIT Press, 1995.
- [3] Fabien Lauer, ChingY. Suen, Gérard Bloch, "A trainable feature extractor for handwritten digit recognition", Elsevier, october 2006.
- [4] Patrice Y. Simard, Dave Steinkraus, John Platt, "Best Practices for Convolutional Noron Networks Applied to Visual Document Analysis," International Conference on Document Analysis and Recognition (ICDAR), IEEE
- [5] Choo Kar Soon, Kueh Chiung Lin, Chung Ying Jeng and Shahrel A. Suandi, "Malaysian Car Number Plate Detection and Recognition System", 2012
- [6] Nobuyuki Otsu, "A threshold selection method from gray-level histograms", 1979

Các website tham khảo:

- [7] <http://noronnetworksanddeeplearning.com/chap6.html>
- [8] <http://noronnetworksanddeeplearning.com/chap2.html>
- [9] www.emgucv.com
- [10] www.opencv.com