

ỨNG DỤNG LÝ THUYẾT XÁC SUẤT THỐNG KÊ TRONG BÀI TOÁN NHẬN DIỆN BIỂN SỐ XE

Nguyễn Thị Mai Trinh
Khoa học và Kỹ thuật thông tin
Trường đại học Công nghệ thông tin
Thành phố Hồ Chí Minh, Việt Nam
21522718@gm.uit.edu.vn

Huỳnh Thị Kim Ngân
Khoa học và Kỹ thuật thông tin
Trường đại học Công nghệ Thông tin
Thành phố Hồ Chí Minh, Việt Nam
21520357@gm.uit.edu.vn

Nguyễn Diệu Phương
Khoa học và Kỹ thuật thông tin
Trường đại học Công nghệ Thông tin
Thành phố Hồ Chí Minh, Việt Nam
215200091@gm.uit.edu.vn

Tóm tắt–Nhận diện biển số xe một bài toán ứng dụng xử lý ảnh nhận được sự quan tâm lớn về hai mặt lý thuyết và thực tiễn. Trong bài này chúng em nghiên cứu mô hình nhận diện biển số xe dựa trên các dữ liệu ảnh mẫu và thực tế. Các thuật toán nhóm sẽ sử dụng là: Lọc ảnh tuyến tính, Histogram, LoG, Convolution, OTSU, Logistic Regression, SVM. Qua báo cáo này cho chúng ta thấy, cách mà mô hình được xây dựng từ toán học và xác suất thống kê.

Từ khóa: Nhận diện biển số xe, Thị giác máy tính, Mean Filte, Histogram, Laplace of Gauss, Convolution, OTSU, Logistic Regression, ZPC

I. Giới thiệu

Bài toán nhận diện biển số xe giúp máy tính từ một ảnh đầu vào có thể nhận ra chính xác biển số xe. Từ đó ứng dụng vào các lĩnh vực: quản lý giao thông, kiểm tra an ninh, quản lý bãi giữ xe,... Đây là một trong những bài toán nhận được sự quan tâm lớn về hai mặt lý thuyết và thực tiễn. Trong đề tài này, nhóm tập trung nghiên cứu các thuật toán ứng dụng xác suất thống kê chuyên sâu trong mô hình nhận diện biển số: nâng cao chất lượng ảnh, nhận diện biển số, phân đoạn ký tự, nhận diện ký tự.

Tuy nhiên, nhóm chỉ mới tìm các tài liệu, báo cáo, nghiên cứu trước và mô phỏng sơ lược trên matlab, chưa có mô hình thực tế.

II. Dữ liệu

Nhóm thu thập dữ liệu từ các bộ data sample công ty GreenParking [1], trang web THỊ GIÁC MÁY TÍNH [2] và ảnh quay chụp thực tế. Hầu hết các ảnh đều khá rõ nét và phù hợp yêu cầu

dữ liệu ban đầu. Tuy nhiên, các ảnh đều theo kích thước biển số cũ, trước khi thông tư 58/2020 được ban hành. Tổng dữ liệu mà nhóm chúng em thu được là 3000 ảnh, bao gồm:

- Biển số xe máy kích thước 280x200
- Biển trước xe ô tô kích thước 470x110
- Biển sau xe ô tô kích thước 190x140

Ngoài ra, để training mô hình nhận diện ký tự trên biển số chúng em sử dụng bộ dữ liệu nổi tiếng về các chữ số viết tay MNIST và bộ ký tự biển số xe máy từ trang web THỊ GIÁC MÁY TÍNH [2]

Việc có nhiều biển số với kích thước, màu sắc, độ sáng khác nhau gây khó khăn trong việc nhận dạng. Để có thể tối ưu hóa và chính xác hơn, chúng em tập trung nghiên cứu mô hình nhận diện biển số Việt Nam. Với các giới hạn:

- Biển hình chữ nhật nền trắng, chữ đen.
- Góc nghiêng biển số xe không quá 45^0
- Hình không bị nhiễu bởi ánh sáng, không bị mờ, ký tự còn phân biệt, nhận diện bằng trực quang được.

III. Phương pháp ứng dụng

A. Lọc ảnh tuyến tính [3]

Dựa vào phép toán áp dụng mà ta có 2 bộ lọc:

- *Bộ lọc tuyến tính*: các điểm ảnh là tuyến tính và đầu ra là mức trung bình của các điểm ảnh lân cận bị mặt nạ lọc bao phủ. Phù hợp với ảnh nhiễu Gauss.
- *Bộ lọc phi tuyến tính*: các điểm ảnh phi tuyến tính và kết quả bộ lọc là hạng của các điểm ảnh lân cận. Phù hợp với ảnh nhiễu muối tiêu Salt & Pepper.

Ở báo cáo này, nhóm sử dụng phương pháp lọc trung bình, là một kỹ thuật lọc tuyến tính hoạt động như một bộ lọc thông thấp.

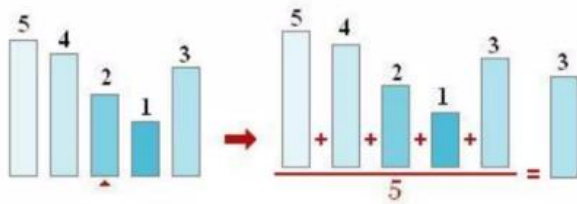
Bước 1: Sử dụng một cửa sổ lọc (ma trận 3x3) quét qua từng điểm của ảnh đầu vào.

Bước 2: Tại mỗi ảnh lấy giá trị của điểm ảnh tương ứng trong vùng 3x3 của ma trận gốc lắp vào ma trận lọc.

Bước 3: Tính giá trị trung bình của tất cả điểm ảnh trong cửa sổ lọc:

$$\{I(q)|q \in W(P)\} \rightarrow AV(P)$$

Bước 4: Gán giá trị trung bình cho ảnh đầu ra



Hình 1. Mô tả bộ lọc trung bình

B. Histogram [4]

Ảnh mức xám (Grayscale Image) là ảnh mà mỗi điểm ảnh (pixel) đại diện cho cường độ ánh sáng tại điểm đó. Bao gồm 256 mức xám khác nhau, giá trị càng cao thì cường độ sáng càng lớn.

Histogram (lược đồ xám) là biểu đồ thống kê tần suất mức sáng xuất hiện trong ảnh.

Các bước cân bằng Histogram:

Bước 1: Tính histogram $P_r(r)$

Bước 2: Xấp xỉ xác suất xuất hiện của mức xám

$$p_r(r_k) = \frac{n_k}{MN}$$

Bước 3: Tính hàm xác suất mật độ

$$s_k = T(r_k) = (L - 1) \sum_{j=0}^k p_r(r_j)$$

Bước 4: Xấp xỉ giá trị mức xám.

Với: r_k : mức xám thứ k của ảnh

k nằm trong khoảng $[0, L-1]$

M, N là chiều dài, chiều rộng của ảnh

n_k : tần suất của mức xám r_k

s_k : cấp độ ánh xạ mức xám của ảnh đầu ra

C. Phát hiện biên Laplace of Gauss (LoG) [5]

Biên là đường bao quanh hoặc đường viền ngăn cách khu vực ảnh liền kề có đặc điểm tương đối khác biệt, một trong số đó là về sự thay đổi đột ngột về cấp xám.

Có 2 phương pháp phát hiện cơ bản:

- **Phát hiện biên gián tiếp:** Phân ảnh thành các vùng ranh giới, giữa các vùng đó là biên

- **Phát hiện biên trực tiếp:** Dựa vào sự biến thiên của về giá trị cấp độ sáng (cấp xám). Chủ yếu dựa vào kỹ thuật đạo hàm. Ở phương pháp này có thể chia thành 2 loại:

- **Phương pháp Gradient:** Ước lượng độ lớn gradient bằng cách sử dụng bộ lọc làm mịn và dự đoán tính toán để xác định vị trí của biên.

- **Phương pháp Laplace:** Tìm kiếm toán tử chéo không tại đạo hàm bậc hai của ảnh để tìm vị trí biên.

Phương pháp dò biên theo toán tử Laplace hiệu quả hơn Gradient trong trường hợp mức xám biến đổi chậm. Tuy nhiên, Laplace rất nhạy với nhiễu, chính vì vậy người ta mở rộng Laplace, dùng hàm Gauss để giảm nhiễu. Phương pháp Laplace of Gauss:

Bước 1: Làm trơn ảnh với bộ lọc Gauss

$$I_{Lap} = I \otimes g(x, y) = g(x, y) \otimes I$$

Trong 1D: $g(x) = e^{-\frac{x^2}{2\sigma^2}}$

Trong 2D: $g(x, y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2 + y^2}{2\sigma^2}}$

Với σ là ngưỡng chuẩn, thể hiện chiều rộng của phân phối Gauss.

Bước 2: Tìm điểm chéo-không

Lấy $|I_{Lap}| = \Delta^2 I_{Lap}$

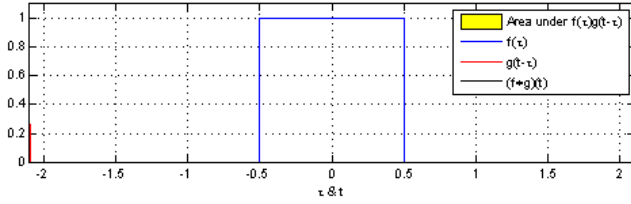
$$\Delta^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

Hiệu chỉnh $I(x, y) = I_{Lap} \geq 0 ? 1 : 0$

D. Phép tính chập (Convolution) [6]

Phép tính chập (Convolution) là toán tuyến tính cho ra kết quả là một hàm dựa trên tính toán 2 hàm đã có. Khác với tương quan chéo, phép chập cần lật kernel theo chiều ngang dọc trước

khi tính tổng của tích. Được ứng dụng trong xác suất, thống kê, thị giác máy tính, xử lý ảnh, xử lý tín hiệu, học máy, và các phương trình vi phân.



Hình 2. Tích chập của 2 xung vuông

Phép tích chập được hình dung thực hiện bằng việc dịch chuyển ma trận kernel (bộ lọc) lần lượt qua tất cả các điểm ảnh trong ảnh, bắt đầu từ góc bên trái trên của ảnh. Và đặt anchor point (điểm neo) tương ứng tại điểm ảnh đang xét. Mỗi lần dịch chuyển, thực hiện tính toán kết quả mới cho điểm ảnh bằng công thức tích chập giữa hàm ảnh $f(x, y)$ và bộ lọc $k(x, y)$ kích thước $m \times n$:

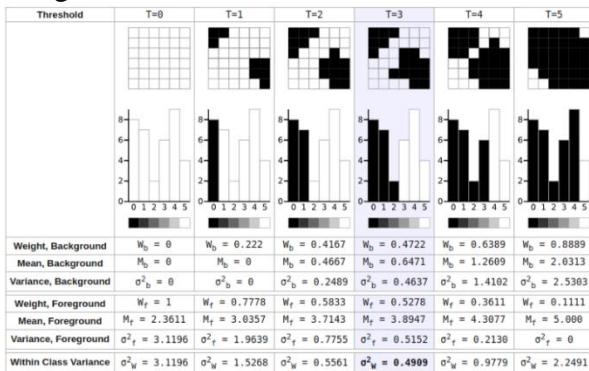
$$\sum_{u=-\frac{m}{2}}^{\frac{m}{2}} \sum_{v=-\frac{n}{2}}^{\frac{n}{2}} k(u, v) f(x - u, y - v)$$

E. Phương pháp phân ngưỡng OTSU [7]

Phương pháp OTSU là một kỹ thuật Thresholding tập trung khai thác và tính toán từ thông tin Histogram của bức ảnh. Bằng việc tính toán trên tất cả các mức Threshold, ta có thể chọn mức thỏa mãn việc phân chia giữa Foreground và Background tốt nhất.

Bước 1: Lần lượt lựa chọn ngưỡng T ban đầu

Bước 2: Tính toán 3 tham số chính: Trọng số, Trung bình, phương sai trên 2 tập background và foreground.



Hình 3. Ví dụ tính toán trên bức ảnh 6x6 với 6 mức ảnh

Bước 3: Tính toán phương sai 2 lớp Within-Class Variance, và chọn ngưỡng có phương sai Within-Class Variance nhỏ nhất.

F. Logistic Regression [8]

Logistic regression (Hồi quy tuyến tính) là một phương pháp thống kê, cho phép phân tích một tập dữ liệu mà trong đó có một hoặc nhiều hơn các biến độc lập có thể tác động tới sự xuất hiện của một kết cục. Logistic Regression cũng giống như Linear Regression nhưng sử dụng hàm sigmoid để dự đoán:

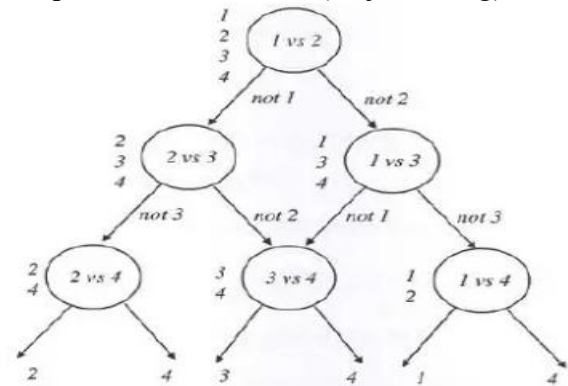
$$P(y|x; w) = \sigma(w^T x)$$

Trong đó $\sigma(-)$ là hàm sigmoid, đầu ra của nó sẽ là khoảng giá trị từ 0 đến 1. Bài toán phân lớp nhị phân tuân theo quy luật Bernoulli, nên xác suất tổng quát trong cả hai trường hợp $\{0, 1\}$ sẽ là:

$$P(y_i|x_i, w) = P(y = 1)^{y_i} (1 - P(y = 1))^{(1-y_i)}$$

Tuy nhiên Logistic Regression chỉ áp dụng được trên bài toán phân lớp nhị phân nên đối với bài toán phân đa lớp thì ta cần sử dụng một trong hai chiến thuật OVO (one-vs-one), OVR (one-vs-rest). Do có nhiều ký tự biểu số có hình dạng giống nhau nên nhóm chọn OVO để tăng độ chính xác phân lớp, chấp nhận hy sinh thời gian.

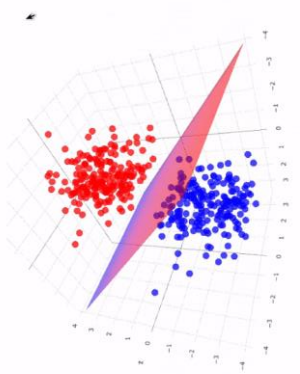
One-vs-one(OVO): Xây dựng rất nhiều bộ phân lớp nhị phân cho từng cặp lớp. Khi có một dữ liệu mới, đưa nó vào toàn bộ các bộ phân lớp nhị phân trên. Kết quả cuối cùng có thể được xác định bằng cách xem lớp nào mà điểm dữ liệu đó được phân vào nhiều nhất (major voting).



Hình 4. Sơ đồ loại trừ OVO

G. SVM (Support Vector Machine) [9]

Phương pháp SVM được thiết kế để giải quyết bài toán phân lớp nhị phân. Ý tưởng của Support Vector Machines là tìm ra siêu phẳng (hyperplane) lớn nhất trong không gian N chiều (N đặc trưng) chia dữ liệu ra làm 2 phần sao cho khoảng cách lề cực đại.



Hình 5. Siêu phẳng có lề cực đại trong không gian 3 chiều

Mỗi điểm trong không gian vector có thể coi là một vector từ gốc tọa độ tới điểm đó. Các điểm dữ liệu nằm trên hoặc gần nhất với siêu phẳng được coi là vector hỗ trợ, ảnh hưởng đến vị trí và hướng của siêu phẳng. Các vector hỗ trợ đều phải cách đều siêu phẳng.

Khoảng cách từ một điểm có tọa độ x_0 tới siêu phẳng có phương trình $w^T x_0 + b = 0$:

$$\frac{|w^T x_0 + b|}{\|w\|_2}$$

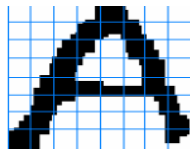
Trong đó: $\|w\|_2 = \sqrt{\sum_{i=1}^d w_i^2}$

d là số chiều không gian

Margin (biên) là khoảng cách gần nhất từ một điểm đến siêu phẳng, việc margin rộng hơn sẽ mang hiệu ứng phân lớp tốt hơn. Chính vì vậy bài toán tối ưu trong SVM chính là đi tìm đường phân chia sao cho margin lớn nhất.

H. Kết hợp các vùng đặc trưng thống kê (ZPC)

- Phân vùng (Zoning):
chia ảnh thành NxN vùng



Hình 6. Trích chọn đặc trưng theo vùng

- Bản đồ chiếu
(Projection histograms)



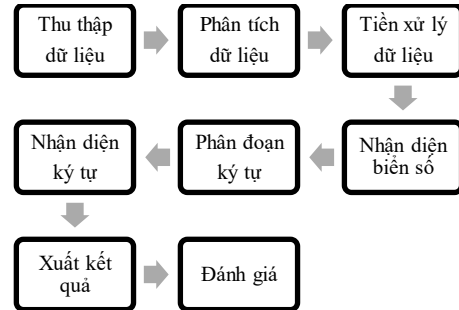
Hình 7. Trích chọn đặc trưng theo biểu đồ chiếu ngang, chiếu dọc và hai đường chéo

- Biên dạng (Contour Profile)



Hình 8. Trích chọn đặc trưng theo biên

IV. Quy trình thực hiện



Hình 9. Quy trình nhận diện biển số cơ bản

Sau đây là quy trình thực hiện đồ án nhận diện biển số xe của nhóm sau khi thu thập và phân tích dữ liệu (đã được trình bày ở mục II)

A. Tiền xử lý dữ liệu

1. Nâng cao chất lượng ảnh

Ngày nay có rất nhiều thuật toán, phương pháp để nâng cao chất lượng ảnh: lọc ảnh, khử nhiễu, làm mờ,... Nhóm em tập trung nghiên cứu và thử nghiệm trên bộ lọc tuyến tính-Lọc trung bình.



Hình 10. Lọc trung bình

2. Biến đổi ảnh R, G, B sang ảnh đa xám

Có 3 phương pháp để chuyển ảnh RGB sang ảnh đa xám (Grayscale)

- Average = $\frac{R+B+G}{3}$
- Lightness = $\frac{\min(R+B+G) + \max(R+B+G)}{2}$
- Luminac = $0.2126*R + 0.7152*G + 0.0722*B$



Hình 11. Chuyển đổi ảnh đa xám

B. Nhận diện biển số



Hình 12. Quy trình nhận diện biển số

Phương pháp lập biểu đồ và phát hiện biên đã được nêu cụ thể ở mục III. Sau đó nhóm nhận diện vùng chứa biển số dựa vào các đặc điểm của biển số xe được nêu ở mục 2.



Hình 13. Phát hiện biên LoG

C. Phân đoạn ký tự



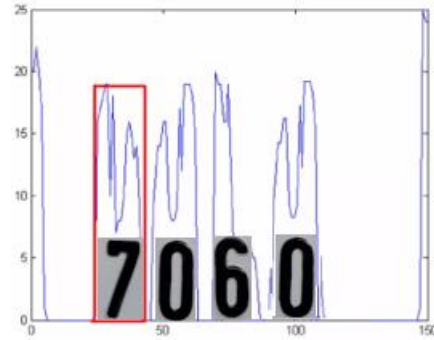
Hình 14. Quy trình phân đoạn ký tự

Để có thể chuyển từ ảnh grayscale sang ảnh nhị phân (chỉ chứa hai giá trị 0,1) cần tìm ra ngưỡng phù hợp. Sau quá trình thử nghiệm phương pháp OTSU để tính toán, nhóm thu được kết quả:

- Với ảnh tối, ít sáng ngưỡng $T = 120$
- Với ảnh nhiều sáng ngưỡng $T = 190$

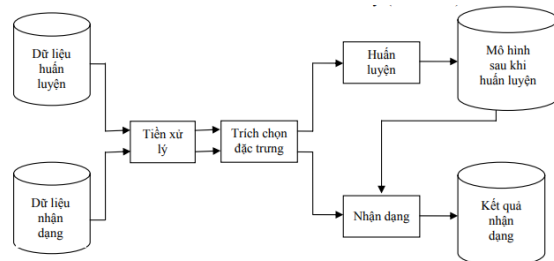
Sau ta chuẩn hóa ảnh bằng cách lấy bù ($0 \rightarrow 1$ và ngược lại). Kết quả, ta thu được nền đen có giá trị 0 và chữ trắng giá trị 1.

Để có thể phân chia thành nhiều ma trận ký tự, ta dựa vào tổng số pixel mức 1-màu trắng. Ở các vùng trống giữa 2 ký tự có tổng pixel rất thấp (giá trị lý tưởng là 0). Ta sử dụng phép chiếu ngang và chiếu dọc để cách ly ký tự.



Hình 15. Tổng số bit theo cột

D. Nhận diện ký tự



Hình 16. Mô hình nhận diện ký tự

Một mô hình nhận diện sẽ có 2 bước chính:

Bước 1: Xây dựng mô hình huấn luyện

Bước 2: Nhận dạng

Dựa vào giá trị các tham số thu được ở 1, một mẫu mới x sau khi qua khâu tiền xử lý và trích chọn đặc trưng, sẽ được đưa qua hàm quét định để xác định lớp của mẫu x .

Ở báo cáo này chúng em trích chọn đặc trưng bằng cách kết hợp: phân vùng, bản đồ chiếu và biên dạng. Tổng đặc trưng mỗi ảnh 16×16 sẽ là 222 đặc trưng. Và sử dụng thuật toán phân lớp SVM áp dụng chiến thuật OVO (One-vs-one)

E. Kết quả và nhận xét

Do thời gian và năng lực có hạn, nên chúng em chỉ mới thử nghiệm trên matlab với từng mô hình riêng rẽ, chưa thể hoàn chỉnh mô hình. Tuy nhiên thông qua quá trình tìm hiểu và thử nghiệm, nhóm nhận thấy:

- Các phương pháp khá dễ hiểu, gần gũi với kiến thức Toán học;
- Mỗi phương pháp đều có ưu nhược điểm rõ ràng.
- Các phương pháp của mô hình chưa thực sự tối ưu và đưa ra kết quả hoàn hảo như mong muốn, phụ thuộc khá nhiều vào chất lượng bộ dữ liệu. Tuy nhiên, có thể dựa vào đó để mở rộng và phát triển hơn trong tương lai.

VI. Kết luận

Mặc dù không đạt được mục đích ban đầu đặt ra, xây dựng một mô hình nhận diện biển số xe hoàn thiện, ứng dụng các kiến thức toán học, xác suất thống kê. Song, qua quá trình tìm tòi nghiên cứu đề tài, chúng em đã tiếp thu thêm nhiều tri thức mới, hiểu sâu hơn cũng như biết cách ứng dụng các kiến thức xác suất thống kê vào mô hình thực tế sẽ như thế nào.

Ngoài ra, lần báo cáo này đã cho chúng em cơ hội để tiếp xúc, làm quen với một trong những đề tài đáng quan tâm trong xã hội hiện nay. Và chúng em tin rằng, hệ thống nhận dạng biển số nói riêng và nhận dạng nói chung sẽ phát triển và ứng dụng rộng rãi trong nhiều lĩnh vực

PHÂN CÔNG NHIỆM VỤ

STT	Nhiệm vụ	Phân công
1	Giao nhiệm vụ và đốc thúc tiến độ	Mai Trinh
2	Slide, video	Huỳnh Ngân
3	Tìm nguồn data, test thuật toán	Diệu Phương
4	Tìm hiểu hạn chế, cách khắc phục	Mai Trinh,
5	Tìm hiểu nội dung các nhóm khác	Huỳnh Ngân

6	Đọc, tìm hiểu tài liệu	Cả nhóm
7	Tìm hiểu thuật toán, xây dựng mô hình	
8	Hoàn thiện, kiểm tra file báo cáo, slide, video	
9	Thuyết trình	

TÀI LIỆU THAM KHẢO

- [1] Bộ ảnh biển số xe máy của công ty GreenParking [trực tuyến], địa chỉ: <https://github.com/thigiacytinh/DataThiGiacMayTinh/blob/main/GreenParking.zip> [truy cập 15/11/2022]
- [2] Thị giác máy tính [trực tuyến], https://thigiacytinh.com/tai-nguyen-xu-ly-anh/tong-hop-data-xu-ly-anh/?fbclid=IwAR2tajA5Ku83kIrb09ovhmb_68Zmdwo9KvV_CSNBCTbuIIsiK_FUM4W4Dh8 [truy cập 15/11/2022]
- [3] Kỹ thuật lọc ảnh và ứng dụng lọc nhiễu làm trơn [trực tuyến], địa chỉ: [Kỹ thuật lọc ảnh và ứng dụng trong lọc nhiễu làm trơn \(slideshare.net\)](http://www.slideshare.net/KiThuatLocAnhVaUngDungTrongLocNhiuLamTron)
- [4] Nguyễn Tiến Đạt, “Tuần 3: Histogram – Histogram equalization” [trực tuyến], địa chỉ: [Tuần 3: Histogram - Histogram equalization \(viblo.asia\)](https://viblo.asia/tuan-3-histogram-histogram-equalization)
- [5] Đặng Thị Thương, “Tìm hiểu phương pháp phát hiện biên cho ảnh đa cấp xám và ảnh màu” [trực tuyến], địa chỉ: [21_DangThiThuong_CT1101.pdf \(hpu.edu.vn\)](http://21.DangThiThuong_CT1101.pdf)
- [6] Phép tính chập trong xử lý ảnh [trực tuyến], địa chỉ: [Phép Tích Chập trong Xử Lý Ảnh \(Convolution\) — Computer Vision \(iostream.vn\)](http://PhapTichChapTrongXuLyAnh(Convolution)-ComputerVision(iostream.vn))
- [7] Tìm hiểu về Otsu threshold [trực tuyến], địa chỉ: [Tìm hiểu về Otsu threshold - Nero's Blog \(nerophung.github.io\)](http://TimHieuVeOtsuThreshold-Nero'sBlog(nerophung.github.io))
- [8] “Bài 10: Logistic Regression” [trực tuyến], địa chỉ: [Machine Learning cơ bản \(machinelearningcoban.com\)](http://MachineLearningCoBan(machinelearningcoban.com))
- [9] Phạm Anh Phương, “Nghiên cứu ứng dụng phương pháp máy véc tơ tựa trong nhận dạng chữ viết tay rời rạc” [trực tuyến], địa chỉ [Microsoft Word - TomTatLuan an - Finish.doc \(itrithuc.vn\)](http://MicrosoftWord-TomTatLuanan-Finish.doc(itrithuc.vn))