

ĐẠI HỌC ĐÀ NẴNG
TRƯỜNG ĐẠI HỌC BÁCH KHOA

TRƯỜNG THANH TÙNG

KIỂM TRA LỖI IN TRÊN REARCASE
SỬ DỤNG MẠNG NƠON NHÂN TẠO

Chuyên ngành: Kỹ thuật Điều khiển và Tự động hóa

Mã số: 60520216

TÓM TẮT LUẬN VĂN THẠC SĨ KỸ THUẬT ĐIỀU KHIỂN
VÀ TỰ ĐỘNG HÓA

Đà Nẵng – Năm 2017

Công trình được hoàn thành tại
TRƯỜNG ĐẠI HỌC BÁCH KHOA

Người hướng dẫn khoa học: **TS. NGUYỄN VĂN MINH TRÍ**

Phản biện 1: TS. Nguyễn Quốc Định

Phản biện 2: TS. Nguyễn Lê Hòa

Luận văn được bảo vệ trước Hội đồng chấm Luận văn tốt nghiệp thạc sĩ Kỹ thuật Điều khiển và Tự động hóa, họp tại Trường Đại học Bách khoa – Đại học Đà Nẵng vào ngày 25 tháng 03 năm 2017.

Có thể tìm hiểu luận văn tại:

- Trung tâm Học liệu, Đại học Đà Nẵng tại trường Đại học Bách khoa.
- Thư viện Khoa Điện, Trường Đại học Bách khoa – Đại học Đà Nẵng.

MỤC LỤC

MỞ ĐẦU	1
NỘI DUNG	4
CHƯƠNG 1: TỔNG QUAN ĐỀ TÀI KIỂM TRA LỖI IN TRÊN REARCASE SỬ DỤNG MẠNG NƠON NHÂN TẠO	4
1.1 Giới thiệu đề tài nghiên cứu	4
1.2 Mục tiêu và nhiệm vụ của đề tài:.....	5
CHƯƠNG 2: CƠ SỞ LÝ THUYẾT VỀ XỬ LÝ ẢNH	6
2.1 Xử lý ảnh, các vấn đề cơ bản trong xử lý ảnh	6
2.2 Xử lý ảnh để nhận dạng kí tự	6
CHƯƠNG 3: NHẬN DẠNG KÝ TỰ DÙNG MẠNG NƠON NHÂN TẠO	7
3.1 Mô tả bài toán.....	7
3.2 Quá trình thực hiện	7
CHƯƠNG 4: CHƯƠNG TRÌNH MÔ PHỎNG VÀ KẾT QUẢ	14
4.1 Môi trường thực nghiệm.....	14
4.2 Tạo cơ sở dữ liệu mẫu	14
4.3 Chương trình chính.....	15
4.4 Hướng phát triển tiếp theo	19
KẾT LUẬN	20

MỞ ĐẦU

LÍ DO CHỌN ĐỀ TÀI

Với sự vận động không ngừng của công nghệ thông tin toàn cầu, mọi tổ chức kinh tế cũng như xã hội, đều mong muốn tận dụng tối đa khả năng cho phép của công nghệ thông tin, để hiện đại hoá quá trình quản lý, sản xuất kinh doanh của mình nhằm đạt hiệu quả kinh tế cao nhất. Nhận dạng là bài toán xuất hiện cách đây khá lâu và vẫn luôn thu hút được nhiều sự quan tâm, nghiên cứu. Đặc biệt là trong vài thập niên gần đây, do sự thúc đẩy của quá trình tin học hoá trong mọi lĩnh vực, bài toán nhận dạng không còn dừng lại ở mức độ nghiên cứu nữa mà nó trở thành một lĩnh vực để áp dụng vào thực tế. Các bài toán nhận dạng đang được ứng dụng trong thực tế hiện nay tập trung vào nhận dạng mẫu, nhận dạng tiếng nói và nhận dạng chữ. Trong số này, nhận dạng chữ là bài toán được quan tâm rất nhiều và cũng đã đạt được nhiều thành tựu rực rỡ. Các ứng dụng có ý nghĩa thực tế lớn có thể kể đến như: nhận dạng chữ in dùng trong quá trình công nghiệp in khắc mã hàng hóa, kiểm tra lỗi tự động, sao lưu sách báo trong thư viện, nhận dạng chữ viết tay dùng trong việc phân loại thư ở bưu điện, thanh toán tiền trong nhà băng và lập thư viện sách cho người mù (ứng dụng này có nghĩa: scan sách bình thường, sau đó cho máy tính nhận dạng và trả về dạng tài liệu mà người mù có thể đọc được). Xuất phát từ yêu cầu của công ty, đang rất cần có những nghiên cứu về vấn đề này. Chính vì vậy em đã chọn đề tài “Kiểm tra

lỗi in trên rearcase sử dụng mạng nơ ron nhân tạo” với mong muốn phân nào áp dụng vào bài toán thực tế ở công ty.

MỤC TIÊU NGHIÊN CỨU

Đề tài tập trung nghiên cứu lí thuyết về công nghệ xử lý ảnh, nhận dạng ký tự sử dụng mạng notron nhân tạo. Sau đó xây dựng ứng dụng xử lý ảnh bằng C#, nhận dữ liệu từ camera, xử lý ảnh, nhận diện kí tự và so sánh với ảnh mẫu để kiểm tra, phát hiện lỗi của khâu sản xuất. Sau đó lưu kết quả vào cơ sở dữ liệu.

PHƯƠNG PHÁP NGHIÊN CỨU

Đề tài được nghiên cứu theo phương pháp kết hợp giữa lý thuyết và nhu cầu thực tế. Cụ thể như sau: - Nghiên cứu các tài liệu liên về công nghệ xử lý ảnh, các vấn đề cơ bản trong xử lý ảnh - Nghiên cứu xử lý ảnh để nhận dạng kí tự - Nghiên cứu nhận dạng kí tự sử dụng mạng notron nhân tạo - Xây dựng ứng dụng bằng C# để đưa vào kiểm thử so với yêu cầu thực tế.

Ý NGHĨA KHOA HỌC VÀ THỰC TIỄN CỦA ĐỀ TÀI

Ý nghĩa khoa học: Triển khai việc ứng dụng công nghệ xử lý ảnh để nhận dạng các ký tự, vật thể tự động, từ đó phân tích các ký tự, vật thể đó theo mục đích của người sử dụng. Ý nghĩa thực tiễn: Tạo ra phần mềm để công việc sản xuất tự

động đạt hiệu quả và có năng suất cao.

CẤU TRÚC LUẬN VĂN

Mở đầu

Chương 1: Tổng quan đề tài nghiên cứu

Giới thiệu về xử lý ảnh, nhận dạng ký tự. Mục tiêu và nhiệm vụ hoàn thành nhận dạng ký tự in trên rearcase.

Chương 2: Cơ sở lý thuyết

Xử lý ảnh và các vấn đề cơ bản trong xử lý ảnh. Xử lý ảnh để nhận dạng ký tự

Chương 3: Nhận dạng ký tự dùng mạng Noron nhân tạo

Mô tả bài toán và quá trình thực hiện

Chương 4: Chương trình mô phỏng và kết quả

Xây dựng chương trình với môi trường thực nghiệm, tạo cơ sở dữ liệu mẫu, chương trình chính. Hướng phát triển tiếp theo.

Kết luận

NỘI DUNG

CHƯƠNG 1:

TỔNG QUAN ĐỀ TÀI KIỂM TRA LỖI IN TRÊN REARCASE SỬ DỤNG MẠNG NORON NHÂN TẠO

1.1 Giới thiệu đề tài nghiên cứu

Đối với mỗi con người chúng ta, việc học và phân biệt được sự khác nhau giữa từng ký tự thật dễ dàng nhưng để máy tính làm được điều tương tự thì không dễ dàng chút nào, việc này cần đến những người có kiến thức tốt về xử lý ảnh mới có thể giúp máy tính học và phân biệt được ký tự nhưng đôi khi, kết quả vẫn không như mong muốn.

Nhận dạng là bài toán xuất hiện cách đây khá lâu và vẫn luôn thu hút được nhiều sự quan tâm, nghiên cứu. Đặc biệt là trong vài thập niên gần đây, do sự thúc đẩy của quá trình tin học hoá trong mọi lĩnh vực, bài toán nhận dạng không còn dừng lại ở mức độ nghiên cứu nữa mà nó trở thành một lĩnh vực để áp dụng vào thực tế. Các bài toán nhận dạng đang được ứng dụng trong thực tế hiện nay tập trung vào nhận dạng mẫu, nhận dạng tiếng nói và nhận dạng chữ. Trong số này, nhận dạng chữ là bài toán được quan tâm rất nhiều và cũng đã đạt được nhiều thành tựu rực rỡ. Các ứng dụng có ý nghĩa thực tế lớn có thể kể đến như: nhận dạng chữ in dùng trong quá trình công nghiệp in khắc mã hàng hóa, kiểm tra lỗi tự động, sao lưu sách báo trong thư viện, nhận dạng chữ viết tay dùng trong việc phân loại thư ở bưu điện, thanh toán tiền trong nhà băng và lập thư viện sách cho người mù (ứng dụng này có nghĩa: scan sách bình

thường, sau đó cho máy tính nhận dạng và trả về dạng tài liệu mà người mù có thể đọc được).

Xuất phát từ yêu cầu thực tế, đang rất cần có những nghiên cứu về vấn đề này. Chính vì vậy em đã chọn đề tài “kiểm tra lỗi in trên rearcase sử dụng mạng nơ ron nhân tạo” với mong muốn phần nào áp dụng vào bài toán thực tế sản xuất.

1.2 Mục tiêu và nhiệm vụ của đề tài:

Nhận thấy tầm quan trọng của công nghệ xử lí ảnh đối với đời sống và sản xuất hiện nay, em đã làm đề tài này với mục tiêu xây dựng ứng dụng xử lí ảnh bằng C#, nhận dữ liệu từ camera, xử lí ảnh, nhận diện kí tự và so sánh với ảnh mẫu để kiểm tra, phát hiện lỗi so với dạng yêu cầu. Sau đó lưu kết quả vào cơ sở dữ liệu.

Nhiệm vụ của chương trình là nhận dạng và so sánh mẫu in với mẫu được yêu cầu là giống nhau. Tương tác với người dùng và tự động báo kết quả.

CHƯƠNG 2:

CƠ SỞ LÝ THUYẾT VỀ XỬ LÝ ẢNH

2.1 Xử lý ảnh, các vấn đề cơ bản trong xử lý ảnh

2.1.1 Xử lý ảnh là gì

2.1.2 Các vấn đề cơ bản trong xử lý ảnh

2.2 Xử lý ảnh để nhận dạng ký tự

2.2.1 Tổng quan

2.2.2 Các phương pháp nhận dạng ký tự

CHƯƠNG 3:

NHẬN DẠNG KÝ TỰ DÙNG MẠNG NƠON NHÂN TẠO

3.1 Mô tả bài toán

Đầu vào : Các dữ liệu đầu vào đã qua xử lý và đạt chuẩn yêu cầu trước khi đưa vào nhận dạng. Các dữ liệu đầu vào là các file ảnh dạng văn bản có định dạng *.bmp, đây là dạng file ảnh chuẩn, có kiểu màu pixel đen và trắng, để việc tiến hành nhận dạng được chính xác. Các file ảnh này đã qua giai đoạn tiền xử lý : xử lý ảnh nhiễu, ảnh đa sắc , lọc ảnh...

Đầu ra mong muốn : là một file text chứa các ký tự Unicode tương ứng mà chương trình mong muốn huấn luyện mạng học thuộc. Số ký tự đầu ra mong muốn phải tương ứng với số kí tự ảnh đầu vào cần huấn luyện.

3.2 Quá trình thực hiện

3.2.1 Xử lí dữ liệu

Lọc nhiễu và làm mịn ảnh. Thường là ảnh thu nhận có nhiều cần phải loại bỏ nhiễu hay ảnh không sắc nét bị mờ hoặc cần làm rõ các chi tiết như đường biên ảnh. Các toán tử không gian dùng trong kỹ thuật tăng cường ảnh được phân nhóm theo công dụng: làm trơn nhiễu, nổi biên.

Nhị phân ảnh: Quá trình phân tích ảnh thực chất là phân tích chuỗi văn bản ảnh thành từng ký tự một. Việc phân tích này dựa trên

việc nhận dạng các pixel màu của ký tự ảnh với dữ liệu đầu vào đã được xử lý và chuyển thành định dạng *.bmp chỉ có 2 pixel điểm đen RGB(0,0,0) và điểm trắng RGB(255, 255, 255).

3.2.2 Tách ký tự

Tách dòng ký tự

Tách từng ký tự

Xác định giới hạn chính xác cho từng ký tự

3.2.3 Ảnh xạ ký tự ảnh vào ma trận giá trị

Sau khi tách, ký tự ảnh sẽ được ánh xạ vào ma trận 2 chiều, có kích thước 15 x 10. Vì các ký tự ảnh dù đã chuẩn hóa, chọn lọc nhưng sau khi tách đều có kích thước khá lớn so với ma trận (khoảng 58x 74), nên các ký tự ảnh này đều phải qua 1 bước chuyển đổi tỉ lệ trước khi chuyển giá trị vào ma trận nhị phân 15 x10.

Việc chuyển đổi tỉ lệ này là một trong những nguyên nhân gây ra sai số dẫn đến nhận dạng ký tự bị sai.

3.2.4 Tìm biên, đường bao

Chương trình sử dụng phương pháp Canny để tìm biên Làm trơn ảnh bằng bộ lọc Gauss nhằm giảm thiểu ảnh hưởng của nhiễu và các chi tiết không mong muốn trong cấu trúc ảnh.

3.2.5 Huấn luyện mạng

a) Xác định cấu trúc mạng

Mạng nơron được xây dựng theo phương pháp học có giám sát. Bài toán lựa chọn mạng Feed-forward 3 lớp với cấu trúc như sau :

Số nơron lớp đầu vào : 150 nơron.

Qua kiểm nghiệm, và dựa trên nhiều tài liệu, nhận thấy giá trị 150 nơron đầu vào là khá phù hợp để nhận dạng với kiểu kí tự ảnh có font chữ trung bình, và phổ biến. Các kí tự ảnh này có chiều cao trung bình là 70 pixel và chiều rộng khoảng 54 pixel. Số nơron lớp ẩn : 500 nơron. Việc lựa chọn giá trị này dựa trên thực nghiệm của các tài liệu chuyên về nhận dạng kí tự quang học. Số nơron này được quyết định từ những kết quả tối ưu và giá trị lỗi cơ sở trong quá trình huấn luyện.

Số nơron đầu ra : 16 nơron tương ứng với 16 bit nhị phân của mã Unicode. Từ 16 nơron này ta tổng hợp và chuyển thành 16 bit nhị phân, sau cùng chuyển thành 1 kí tự Unicode.

b) Thuật toán huấn luyện mạng.

Mạng feed- forward sử dụng giải thuật lan truyền ngược sai số Back Propagation. Giải thuật gồm 2 bước :

- Bước 1: Lan truyền xuôi đầu vào qua mạng.

Sử dụng các công thức :

Công thức chung tính đầu ra của một nơron thứ i tại lớp thứ k :

$$y_i = f \left(\sum_{j=0}^n w_{ij} x_j - b_i \right)$$

Với f là hàm chuyển sigmoid lưỡng cực được tính theo công thức:

$$f = \frac{2}{1 + e^{-\alpha t}} - 1$$

α : hệ số góc của hàm chuyển

t : biến net-input

b_i : hệ số ngưỡng hay độ lệch

Áp dụng đối với mô hình mạng của chương trình : Công thức cho đầu ra của một nơron thứ i ($1 \rightarrow 250$) tại lớp ẩn :

$$a_i = f \left(\sum_{j=1}^n w_{ij} x_j - b_i \right)$$

Với w_{ij} : trọng số tại nơron thứ i của lớp ẩn kết nối với đầu vào thứ j của lớp vào

x_j : giá trị đầu vào của nơron thứ j tại lớp vào

b_i : giá trị ngưỡng hay độ lệch của nơron thứ i của đầu vào

Công thức cho đầu ra của một nơron thứ k ($k = 1 \rightarrow 16$) tại lớp output:

$$y_k = f \left(\sum_{j=1}^n w_{kj} a_j - b_k \right) = f \left[\sum_{j=1}^n w_{kj} \left(f \left(\sum_{j=1}^n w_{ij} x_j \right) \right) \right]$$

- Bước 2 : Lan truyền ngược

Tính toán sai lệch giữa đầu ra thực và đầu ra mong muốn của neuron thứ k tại đầu ra $e_i = t_i - y_i$

Tổng bình phương sai số của mạng ứng với mẫu học (Xs, Ts):

$$E = \frac{1}{2} \sum_{k=1}^p (t_i - y_i)^2$$

P: số phần tử đầu ra. Nếu $E < \varepsilon$ và $l+1 < \text{epochs}$ thì :Thông tin sai số sẽ được lan truyền ngược qua mạng để điều chỉnh lại trọng số tại vòng lặp l.Công thức điều chỉnh trọng số với liên kết giữa neuron thứ j trong lớp ẩn và neuron thứ i trong lớp ra tại lần lặp l+1: ($l+1 < \text{số lần dạy(epochs)}$)

$$w_{ij}(l+1) = w_{ij}(l) + \eta \cdot e_i(l) \cdot y_j \cdot f'(y_i(l))$$

f' : đạo hàm của hàm chuyển lưỡng cực ,công thức :

$$f' = \frac{1 - x^2}{2}$$

Công thức điều chỉnh trọng số với liên kết giữa neuron vào thứ j và neuron ẩn thứ i , tại lần lặp thứ l+1 ($l+1 < \text{epochs}$)

$$w_{ij}(l+1) = w_{ij}(l) + \eta \cdot x_j \cdot f'(y_i) \cdot \sum_{k=1}^m w_{ki}(l+1) \cdot e_k(l) \cdot f'(y_k(l))$$

Các tham số sử dụng trong chương trình

Tốc độ học $\eta = 150$.

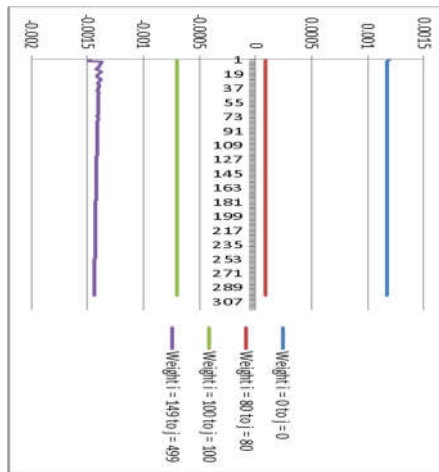
Hệ số góc α Sigmoid= 0.014.

Giá trị ngưỡng hay độ lệch : 30

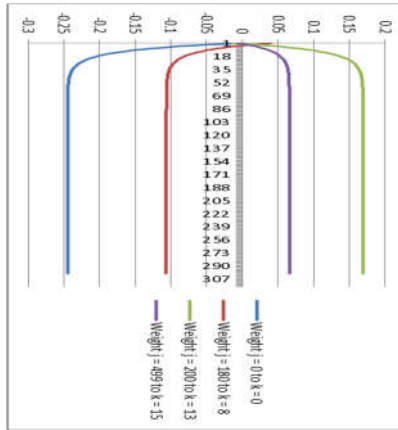
Số lần dạy 300-600 tùy độ phức tạp của từng loại font.

Ngưỡng của lỗi = 0.0002.

Khảo sát sự hội tụ của các trọng số



Hình 3.9 Đồ thị sự biến thiên của W_{ij}



Hình 3.10 Đồ thị sự biến thiên của W_{jk}

CHƯƠNG 4:

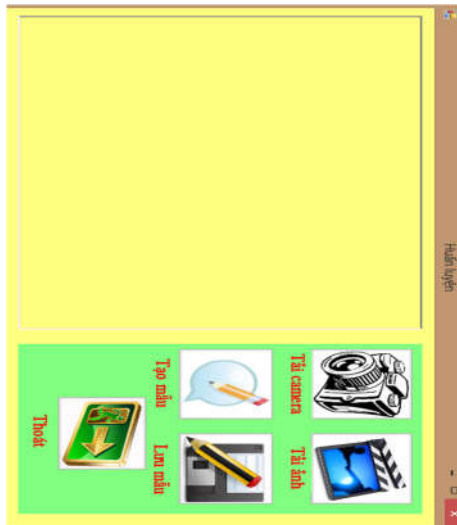
CHƯƠNG TRÌNH MÔ PHỎNG VÀ KẾT QUẢ

4.1 Môi trường thực nghiệm

Chương trình được cài đặt trên ngôn ngữ Visual C# và được thử nghiệm trên hệ điều hành Windows XP SP2, máy tính PC tốc độ 1,6 GHz, bộ nhớ 256MB RAM.

4.2 Tạo cơ sở dữ liệu mẫu

Trước tiên chương trình cần học qua bộ dữ liệu mẫu. Trong quá trình sử dụng, nếu người dùng cần định nghĩa một ký tự mới thì chương trình có thêm chức năng cho người dùng tự định nghĩa.



Hình 4. 1: Giao diện chương trình tạo cơ sở dữ liệu phục vụ cho việc nhận dạng

4.3 Chương trình chính

4.3.1 Đăng nhập

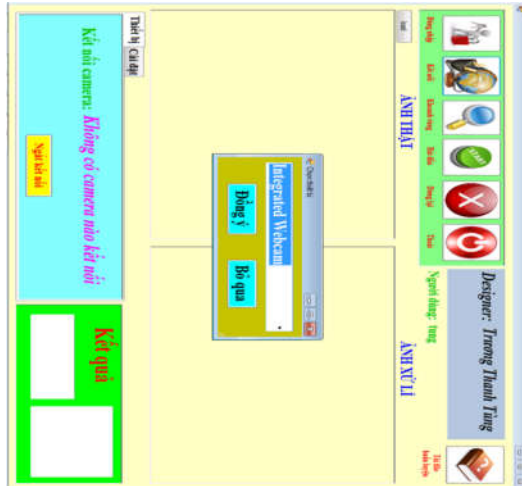
Để sử dụng các chức năng có trong chương trình thì đầu tiên ta phải đăng nhập (ở đây mặc định là user và password như hình vẽ).



Hình 4.2: giao diện chương trình đăng nhập

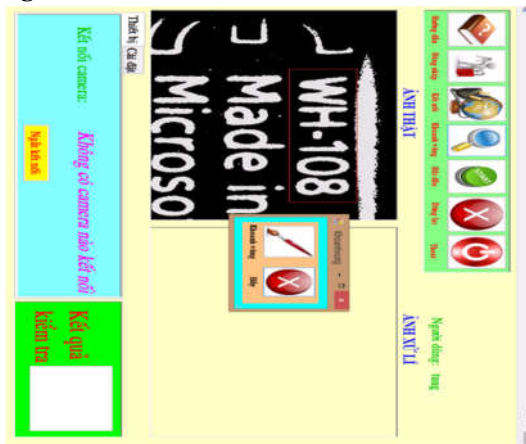
4.3.2 Chọn camera

Tiếp theo sau khi đăng nhập ta sẽ tiến hành kết nối với camera



Hình 4.3 chọn camera kết nối

4.3.3 Khoanh vùng

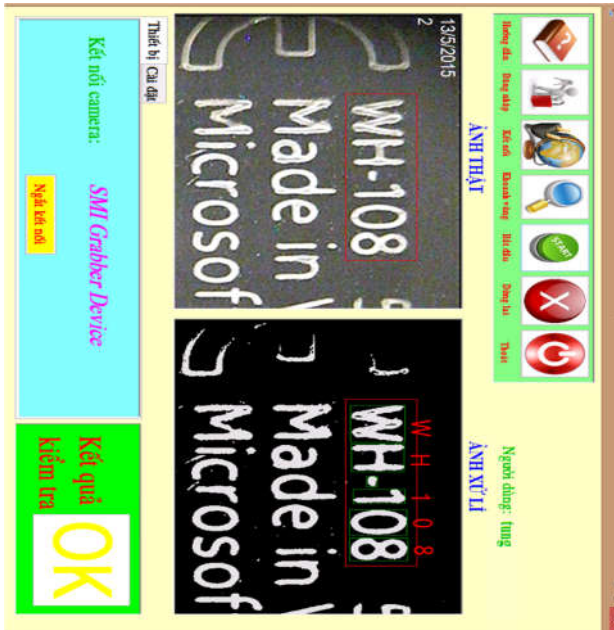


Hình 4.4 khoanh vùng kí tự

4.3.4 Chạy chương trình

Ấn bắt đầu để chạy chương trình. Cứ sau 5s sẽ xử lí một lần và hiện kết quả xem con hàng có được in đúng hay không ở ô “ Kết quả kiểm tra”.

4.3.5 Kết quả chương trình



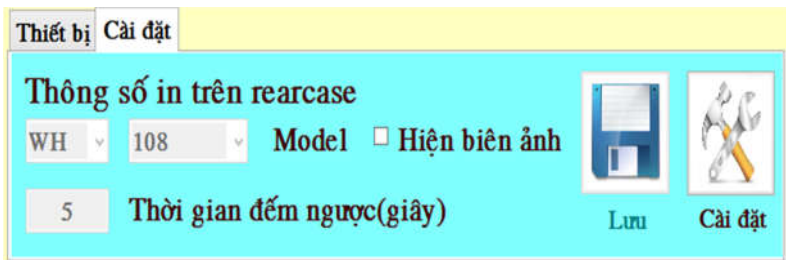
Hình 4. 5: Kết quả với mã hàng WH-108



Hình 4. 6: Kết quả với mã hàng WH-109

4.3.6 Một số chức năng khác

Ngoài ra còn một số cài đặt khác như chọn mã hàng kiểm tra hay chọn thời gian xử lí hoặc hiện đường biên ảnh đã xử lí trong chương trình:



Hình 4. 7: Một số cài đặt khác

4.4 Hướng phát triển tiếp theo

Từ quá trình thực nghiệm cho thấy, trương trình đã có những thành công nhất định. Song bên cạnh đó cũng còn khá nhiều nhược điểm cần cải tiến. Trong quá trình hoàn thành đồ án tôi nhận thấy mình khá tâm huyết với bài toán này và nhận thấy đây là bài toán có khả năng phát triển cao. Tôi rất mong muốn sẽ phát triển nó thành một trương trình có thể ứng dụng trong thực tế. Tôi xin đưa ra một số hướng phát triển tiếp theo cho bài toán

- Phát triển để trương trình có thể tích hợp với các trương trình xử lý tự động của Robot. Phát triển thêm về mặt dữ liệu để trương trình có thể hoạt động tốt với dữ liệu của nhiều loại ký tự đầu vào hơn.
- Cải tiến một số thuật toán tiền xử lý để chương trình có tốc độ tốt hơn.
- Phát triển giao diện thân thiện hơn với người sử dụng.

KẾT LUẬN

Đề tài nghiên cứu về những thuật toán trong xử lý ảnh sau đó áp dụng vào việc nhận dạng ký tự in của rearcase trong sản xuất tự động. Với yêu cầu nhận dạng được ký tự in trên rearcase từ hình chụp đầu vào và hiển thị dưới dạng văn bản thì chương trình sau khi xây dựng đã đáp ứng tốt yêu cầu đó. Điều đó cho thấy chương trình có thể được sử dụng để xây dựng hệ thống nhận dạng ký tự in với những ứng dụng thực tế như hệ thống nhận dạng ký tự và kiểm tra lỗi sau khi in, ... So với những chương trình nhận dạng trước đây, đề tài có những thay đổi đáng kể trong phương pháp nhận dạng và ngôn ngữ sử dụng là C#, không như các chương trình trước sử dụng c++ hay matlab.

Qua thời gian làm đề tài tốt nghiệp, tác giả đã xây dựng được chương trình mô phỏng nhận dạng ký tự dựa trên ngôn ngữ C# và ứng dụng bài toán vào thực tế sản xuất rearcase và yêu cầu của đối tác ở môi trường công nghiệp.

Tuy nhiên, với khoảng thời gian cho phép và kiến thức có hạn, đề tài không thể tránh khỏi những vấn đề chưa giải quyết được, tiêu biểu nhất là việc tìm kiếm một tập mẫu có kích thước lớn và hoàn chỉnh để đáp ứng được yêu cầu của đề tài vẫn rất khó khăn, số lượng tập tin mẫu có kích thước nhỏ dẫn đến có sự sai lệch khi nhận dạng một số ký tự có hình dạng giống nhau (số 0 với chữ D hay số 8 với chữ B, ...). Bên cạnh đó, ảnh rearcase được chụp trong những điều kiện ánh sáng khác nhau cũng ảnh hưởng lớn đến kết quả của chương trình.