

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

TRỊNH THỊ PHƯƠNG THẢO

**NHẬN DẠNG CHỮ VIẾT TAY TIẾNG VIỆT
SỬ DỤNG MẠNG NƠRON**

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

Thái Nguyên, Tháng 06 năm 2012

LỜI CẢM ƠN

Xin chân thành cảm ơn Thầy giáo, PGS.TS. **Lê Bá Dũng** đã tận tình chỉ dạy, hướng dẫn tôi trong suốt thời gian nghiên cứu và thực hiện luận văn.

Tôi cũng xin chân thành cảm ơn các Thầy giáo Viện Công nghệ Thông tin và các Thầy giáo Trường Công nghệ thông tin và Truyền thông - ĐH Thái Nguyên đã giảng dạy, giúp đỡ trong suốt thời gian học tập.

Xin cảm ơn tất cả các anh chị học viên Cao học khóa 9, cảm ơn các cán bộ công chức, giảng viên Trường Công nghệ thông tin và Truyền thông - ĐH Thái Nguyên đã tạo điều kiện tốt cho tôi trong suốt trong hai năm học qua.

Xin cảm ơn các bạn bè, đồng nghiệp đã tạo điều kiện cũng như đã chỉ bảo tôi rất nhiều trong thời gian thực hiện luận văn này.

Cuối cùng, xin chân thành cảm ơn gia đình đã động viên và tạo mọi điều kiện thuận lợi để tôi có được kết quả như ngày hôm nay.

Thái Nguyên, tháng 06/2012

Người viết luận văn

Trịnh Thị Phương Thảo

LỜI CAM ĐOAN

Tôi xin cam đoan đề tài luận văn “**Nhận dạng chữ viết tay Tiếng Việt sử dụng mạng Nơron**” là công trình nghiên cứu của bản thân tôi. Các số liệu, kết quả nghiên cứu nêu trong luận văn này là trung thực và không sao chép y nguyên từ một công trình nào khác. Tôi xin chịu trách nhiệm về luận văn của mình.

Thái Nguyên, tháng 06/2012

Người viết luận văn

Trịnh Thị Phương Thảo

MỤC LỤC

	Trang
TRANG PHỤ BÌA.....	
LỜI CẢM ƠN	i
LỜI CAM ĐOAN	ii
MỤC LỤC.....	iii
DANH MỤC CÁC TỪ TIẾNG ANH VIẾT TẮT.....	vii
DANH MỤC CÁC HÌNH.....	viii
PHẦN MỞ ĐẦU.....	1
CHƯƠNG I	3
TỔNG QUAN VỀ MẠNG NƠN NHÂN TẠO VÀ MẠNG NƠN	
KOHONEN	3
1.1. Giới thiệu về mạng nơron nhân tạo	3
1.1.1 Lịch sử phát triển.....	3
1.1.2. Mô hình mạng nơ-ron nhân tạo.....	4
1.1.2.1. Nơ-ron sinh học	4
1.1.2.2. Nơ-ron nhân tạo	6
1.1.2.3. Mạng nơ-ron nhân tạo.....	12
1.1.2.3.1. Mạng truyền thẳng	13
1.1.2.3.2. Mạng hồi quy (Recurrent Neutral Network).....	15
1.1.2.4. Luật học	16
1.2. Phạm vi ứng dụng của mạng nơron nhân tạo	20
1.2.1. Những bài toán thích hợp.....	20
1.2.1.1. Phân loại	21
1.2.1.2. Mô hình hoá.....	21

1.2.1.3. Biến đổi.....	23
1.2.1.4. Liên kết	23
1.2.2. Các lĩnh vực ứng dụng mạng nơron.....	24
1.3. Mạng KOHONEN.....	26
1.3.1. Giới thiệu về mạng Nơron Kohonen.....	26
1.3.2. Cấu trúc của mạng nơron Kohonen [9].....	27
1.3.3. Thực hiện mạng nơron Kohonen [9].....	28
1.3.3.1 Chuẩn hóa đầu vào.....	28
1.3.3.2 Tính toán đầu ra cho mỗi nơron.....	29
1.3.3.3 Chọn nơron chiến thắng.....	29
1.3.3.4 Quá trình học của mạng nơron Kohonen.....	30
1.3.3.4.1 Tốc độ học	32
1.3.3.4.2 Hiệu chỉnh trọng số	33
1.3.3.4.3 Tính toán lỗi	34
CHƯƠNG II.....	36
ỨNG DỤNG MẠNG NƠRON KOHONEN TRONG BÀI TOÁN NHẬN DẠNG CHỮ VIẾT TAY TIẾNG VIỆT.....	36
2.1. Tổng quan về bài toán nhận dạng.....	36
2.2. Giới thiệu về bài toán nhận dạng chữ viết tay Tiếng Việt.....	39
2.3. Phương pháp nhận dạng chữ viết tay sử dụng mạng nơron	43
2.4. Phát biểu bài toán.....	44
2.5 Các bước giải quyết bài toán sử dụng mạng nơron Kohonen	44
2.5.1. Xây dựng giao diện vẽ	44
2.5.2. Xây dựng mạng nơron Kohonen.....	47
2.5.3. Xử lý dữ liệu (phân tích ảnh)	47

2.5.4. Huấn luyện mạng nơron Kohonen	48
2.5.5 Nhận dạng mạng nơron Kohonen	50
2.5.6 Kết luận	50
CHƯƠNG III.....	52
MÔ PHỎNG NHẬN DẠNG KÝ TỰ VIẾT TAY TIẾNG VIỆT RỜI RẠC SỬ DỤNG MẠNG NƠRON KOHONEN.....	52
3.1 Giới thiệu	52
3.2 Thực hiện chương trình nhận dạng ký tự	53
3.2.1 Xác định các tham số cho mạng.....	54
3.2.2 Vẽ hình ảnh	55
3.2.3 Lấy mẫu xuống hình ảnh.....	56
3.2.3.1 Lưu hình ảnh được lấy mẫu xuống.....	56
3.2.3.2 Kích thước và vị trí.....	56
3.2.3.3 Thực hiện lấy mẫu xuống	57
3.2.4 Các tập huấn luyện	58
3.2.5 Lớp mạng cơ bản.....	58
3.2.6 Lớp mạng nơron Kohonen	59
3.2.6.1 Tính toán đầu ra của mạng Kohonen.....	61
3.2.6.2 Huấn luyện mạng nơron Kohonen.....	62
3.3 Chương trình mô phỏng.....	65
3.3.1 Các chức năng của chương trình	65
3.3.2 Kết quả nhận dạng.....	66
3.3.3 Kết luận	68

CHƯƠNG IV: KẾT LUẬN	71
4.1 Các kết quả đã đạt được.....	71
4.1.1 Về mặt lý thuyết	71
4.1.2 Về mặt thực tiễn	72
4.2 Hướng phát triển tiếp theo	73
4.3 Một vài suy nghĩ sau khi nghiên cứu.....	73
TÀI LIỆU THAM KHẢO.....	75

DANH MỤC CÁC TỪ TIẾNG ANH VIẾT TẮT

STT	Từ viết tắt	Từ viết tương minh
1	HMM	Hidden Markov Model
2	LVQ	Learning Vector Quantization
3	MLP	Multi Layer Perception
4	OCR	Optical Character Recognition
5	PDA	Personal Digital Assistant
6	RFID	Radio Frequency Identification
7	SOM	Self Organizing Maps
8	SVM	Support Vector Machine

DANH MỤC CÁC HÌNH

Hình 1.1. Mô hình nơ-ron sinh học.....	6
Đồ thị hàm đồng nhất (Identity function)	8
Đồ thị hàm bước nhị phân (Binary step function)	9
Đồ thị hàm sigmoid.....	10
Đồ thị hàm sigmoid lưỡng cực.....	10
Hình 1.2. Mô hình một nơ-ron	11
Hình 1.3. Mạng truyền thẳng một lớp.....	14
Hình 1.4. Mạng truyền thẳng nhiều lớp	15
Hình 1.5. Mạng một lớp có nối ngược.....	16
Hình 1.6. Mạng nhiều lớp có nối ngược	16
Hình 1.7: Cấu trúc của mạng Kohonen.....	27
Hình 1.8: Mạng nơron Kohonen điển hình	28
Hình 1.9: Huấn luyện mạng Kohonen (SOM)	30
Hình 1.10. Quá trình huấn luyện mạng nơron Kohonen.....	31
Hình 2.1: Mô hình chung trong nhận dạng chữ viết.....	42
Hình 2.2: Quá trình tìm giới hạn ký tự.....	45
Hình 2.3: Quá trình lấy mẫu xuống.....	45
Hình 2.4 Quá trình ánh xạ từ ma trận điểm sang ma trận giá trị	48
Hình 2.5 Quá trình huấn luyện mạng nơron Kohonen.....	49
Hình 3.1 Giao diện chương trình mô phỏng	65
Hình 3.2 Nhận dạng ký tự O	67
Hình 3.3 Nhận dạng ký tự Ô	67
Hình 3.4 Nhận dạng ký tự U'	68

