

**ĐẠI HỌC THÁI NGUYÊN**  
**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN & TRUYỀN THÔNG**

**MAI VĂN THỦY**

**NGHIÊN CỨU VỀ MÔ HÌNH THỐNG KÊ HỌC SÂU  
VÀ ỨNG DỤNG TRONG NHẬN DẠNG CHỮ VIẾT  
TAY HẠN CHẾ**

**LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH**

**THÁI NGUYÊN - 2015**

**ĐẠI HỌC THÁI NGUYÊN**  
**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN & TRUYỀN THÔNG**

**MAI VĂN THỦY**

**NGHIÊN CỨU VỀ MÔ HÌNH THỐNG KÊ HỌC SÂU  
VÀ ỨNG DỤNG TRONG NHẬN DẠNG CHỮ VIẾT  
TAY HẠN CHẾ**

***Chuyên ngành* : Khoa Học Máy Tính**  
***Mã số* : 60 48 01**

**LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH**

**NGƯỜI HƯỚNG DẪN KHOA HỌC**  
**TS. Vũ Tất Thắng**

**THÁI NGUYÊN - 2015**

## LỜI CAM ĐOAN

Tôi cam đoan đây là công trình nghiên cứu của riêng tôi.

Các số liệu, kết quả nêu trong luận văn là trung thực và chưa từng được ai công bố trong bất kỳ công trình nào khác.

Qua đây tôi xin chân thành cảm ơn toàn thể các thầy cô trong khoa đào tạo sau đại học Trường Đại học Công nghệ Thông tin và Truyền thông – Đại học Thái Nguyên, những người đã trực tiếp giảng dạy, truyền đạt cho tôi kiến thức chuyên môn và phương pháp làm việc khoa học.

Đặc biệt, tôi xin chân thành cảm ơn **TS. Vũ Tất Thắng**, đã tận tình hướng dẫn để tôi có thể hoàn thành luận văn này.

Tôi cũng xin gửi lời cảm ơn tới gia đình bạn bè đã giúp đỡ, động viên và tạo điều kiện cho tôi trong quá trình làm luận văn.

**Tác giả luận văn**

*Mai Văn Thủy*

## MỤC LỤC

<b>LỜI CAM ĐOAN</b> .....	<b>i</b>
<b>MỤC LỤC</b> .....	<b>iv</b>
<b>DANH MỤC HÌNH ẢNH</b> .....	<b>vi</b>
<b>DANH MỤC BẢNG BIỂU</b> .....	<b>vii</b>
<b>LỜI MỞ ĐẦU</b> .....	<b>1</b>
<b>Chương 1: GIỚI THIỆU ĐỀ TÀI</b> .....	<b>3</b>
<b>1.1. Giới thiệu về bài toán nhận dạng</b> .....	<b>3</b>
1.1.1. Các giai đoạn phát triển .....	3
1.1.2. Tình hình nghiên cứu trong nước .....	4
1.1.3. Tình hình nghiên cứu ở nước ngoài.....	4
<b>1.2. Các bước xử lý cho bài toán nhận dạng hoàn chỉnh</b> .....	<b>6</b>
<b>1.3. Kết luận chương</b> .....	<b>8</b>
<b>Chương 2: MÔ HÌNH SVM VÀ MÔ HÌNH THỐNG KÊ HỌC SÂU</b> .....	<b>9</b>
<b>2.1. Tổng quan về mô hình SVM (Support Vector Machine)</b> .....	<b>9</b>
2.1.1. Cơ sở lý thuyết.....	9
2.1.1.1. Giới thiệu bài toán phân lớp nhị phân.....	9
2.1.1.2. Máy SVM tuyến tính.....	10
2.1.1.3. Máy SVM phi tuyến.....	17
2.1.2. Các thuật toán huấn luyện SVM .....	19
2.1.2.1. Thuật toán chặt khúc .....	19
2.1.2.2. Thuật toán phân rã.....	19
2.1.2.3. Thuật toán cực tiểu tuần tự.....	20
<b>2.2. Cơ sở lý thuyết mô hình thống kê học sâu</b> .....	<b>23</b>
2.2.1. Một số lý thuyết về mạng Neuron .....	23
2.2.1.1. Giới thiệu về mạng Neuron.....	23
2.2.1.2. Cấu trúc và hoạt động của mạng Neuron.....	23
2.2.1.3. Quá trình huấn luyện mạng và các thuật toán học mạng .....	28

2.2.2. Hopfield Network .....	31
2.2.2.1. Cấu trúc mạng Hopfield.....	31
2.2.2.2. Mạng Hopfield rời rạc.....	33
2.2.2.3. Mạng Hopfield liên tục .....	34
2.2.3. Boltzmann Machines .....	36
2.2.4. Restricted Boltzmann Machines .....	39
2.2.3. Thuật toán lan truyền ngược .....	40
<b>2.3. Kết luận chương .....</b>	<b>42</b>
<b>Chương 3: KẾT QUẢ THỰC NGHIỆM VÀ ĐÁNH GIÁ .....</b>	<b>43</b>
<b>3.1. Môi trường thực nghiệm .....</b>	<b>43</b>
<b>3.2. Dữ liệu thực nghiệm .....</b>	<b>43</b>
<b>3.3. Kết quả thực nghiệm với mô hình SVM.....</b>	<b>44</b>
<b>3.4. Huấn luyện mô hình và kết quả thực nghiệm với mô hình thống kê học sâu.....</b>	<b>45</b>
3.4.1. Huấn luyện mô hình.....	45
3.4.1.1. Cấu trúc mô hình.....	45
3.4.1.2. Phương pháp huấn luyện và quy trình nhận dạng.....	47
3.4.2. Giao diện chính của chương trình.....	48
3.4.3. Kết quả thực nghiệm.....	49
<b>3.5. Đánh giá kết quả thực nghiệm của hai mô hình.....</b>	<b>50</b>
<b>3.6. Kết luận chương .....</b>	<b>51</b>
<b>KẾT LUẬN CHUNG .....</b>	<b>52</b>
<b>TÀI LIỆU THAM KHẢO .....</b>	<b>54</b>
<b>PHỤ LỤC: HUẤN LUYỆN MÔ HÌNH .....</b>	<b>56</b>

## DANH MỤC HÌNH ẢNH

<i>Hình 1-1:</i> Các bước trong nhận dạng chữ viết tay .....	6
<i>Hình 2-1:</i> Các siêu phẳng $H_1, H_2$ phân cách giữa hai lớp .....	9
<i>Hình 2-2:</i> Siêu phẳng tách tuyến tính .....	10
<i>Hình 2-3:</i> Không thể phân hoạch tập mẫu trên bằng một siêu phẳng.....	13
<i>Hình 2-4:</i> Một mặt phân chia phi tuyến có thể trở thành một siêu phẳng trong không gian lớn hơn. ....	17
<i>Hình 2-5:</i> Cấu trúc của một neuron .....	24
<i>Hình 2-6:</i> Cấu trúc chung của mạng neuron .....	26
<i>Hình 2-7:</i> Cấu trúc của mạng Hopfield .....	31
<i>Hình 2-8:</i> Đồ thị hàm satlins .....	32
<i>Hình 2-9:</i> Mạng Hopfield liên tục sử dụng mạch điện tử. ....	35
<i>Hình 2-10:</i> Một Boltzmann Machine với 3 nút ẩn.....	36
<i>Hình 2-11:</i> Một RBM đơn giản với 3 hidden units và 2 visible units.....	39
<i>Hình 3-2:</i> Giao diện chính của chương trình nhận dạng chữ viết tay hạn chế .....	48
<i>Hình 3-3:</i> Chương trình khi nhận dạng 1 ảnh bất kỳ .....	48
<i>Hình 3-4:</i> Nhận dạng và thống kê nhiều ảnh .....	49

**DANH MỤC BẢNG BIỂU**

<i>Bảng 2-1:</i> Các hàm truyền cơ bản.....	27
<i>Bảng 3-1:</i> Kết quả thực nghiệm mô hình SVM trên 2 tập dữ liệu MNIST và Tuyển Sinh.....	44
<i>Bảng 3-2:</i> Kết quả thực nghiệm mô hình SVM trên tập dữ liệu MNIST .....	44
<i>Bảng 3-3:</i> Kết quả thực nghiệm mô hình SVM trên tập dữ liệu Tuyển Sinh .....	45
<i>Bảng 3-4:</i> Kết quả thực nghiệm trên 2 tập dữ liệu MNIST và Tuyển Sinh.....	49
<i>Bảng 3-5:</i> Kết quả thực nghiệm trên tập dữ liệu Tuyển sinh.....	50
<i>Bảng 3-6:</i> Bảng so sánh kết quả giữa hai mô hình .....	50

## LỜI MỞ ĐẦU

Nhận dạng chữ viết tay là bài toán khó trong lớp các bài toán nhận dạng chữ, và vẫn luôn thu hút được nhiều sự quan tâm nghiên cứu của các nhà khoa học. Đặc biệt là trong vài thập niên gần đây, do sự thúc đẩy của quá trình tin học hóa trong mọi lĩnh vực, ứng dụng nhận dạng chữ càng có nhiều ý nghĩa khi được sử dụng cho các bài toán trong thực tế. Cũng như nhiều bài toán nhận dạng tiếng nói, hình ảnh... khác, thì độ chính xác của hệ thống vẫn tiếp tục cần phải cải thiện nhằm vươn tới khả năng nhận dạng giống như con người.

Tuy nhiên, với bài toán nhận dạng chữ viết tay thì vấn đề trở nên phức tạp hơn nhiều so với bài toán nhận dạng chữ in thông thường ở những vấn đề sau đây [3]:

- Với chữ viết tay thì không thể có các khái niệm font chữ, kích cỡ chữ. Các kí tự trong một văn bản chữ viết tay thường có kích thước khác nhau. Thậm chí, cùng một kí tự trong một văn bản do một người viết nhiều khi cũng có độ rộng, hẹp, cao, thấp khác nhau,...
- Với những người viết khác nhau chữ viết có độ nghiêng khác nhau (chữ nghiêng nhiều/ít, chữ nghiêng trái/phải...).
- Các kí tự của một từ trên văn bản chữ viết tay đối với hầu hết người viết thường bị dính nhau vì vậy rất khó xác định được phân cách giữa chúng.
- Các văn bản chữ viết tay còn có thể có trường hợp dính dòng (dòng dưới bị dính hoặc chồng lên dòng trên).

Trong những năm gần đây, mô hình mạng Neuron theo hướng học sâu đã cho thấy những kết quả tốt trong nhiều bài toán khác nhau, trong đó có nhận dạng chữ.

Xuất phát từ yêu cầu thực tế, đang rất cần có nhưng nghiên cứu về vấn đề này. Chính vì vậy học viên đã chọn đề tài “***Nghiên cứu về mô hình thống kê học sâu và ứng dụng trong nhận dạng chữ viết tay hạn chế***” làm luận văn tốt nghiệp với mong muốn phần nào áp dụng vào bài toán thực tế.

Bài toán đã đặt ra phải giải quyết được những yêu cầu sau:

- ✓ Nhận dạng được các ký tự từ ảnh đầu vào



- ✓ Tiến hành nhận dạng kí tự đơn lẻ sử dụng mạng Neuron nhân tạo theo phương pháp học sâu Restricted Boltzmann machine (RBM).
- ✓ Đánh giá kết quả và so sánh với mô hình Support Vector Machine

Với những yêu cầu đã đặt ra ở trên, cấu trúc của luận văn sẽ bao gồm những nội dung sau đây:

➤ **Chương 1: Tổng quan về đề tài**

Giới thiệu về bài toán nhận dạng chữ viết tay, tình hình nghiên cứu trong và ngoài nước, quy trình chung để giải quyết bài toán và các phương pháp điển hình trong việc huấn luyện nhận dạng, phạm vi của đề tài.

➤ **Chương 2: Mô hình SVM và mô hình thống kê học sâu**

Trình bày về cơ sở lý thuyết của mô hình SVM (Support Vector Machine) và huấn luyện trong bài toán nhận dạng chữ viết tay. Cơ sở lý thuyết của mô hình thống kê học sâu: Hopfield network, Boltzmann Machines, Restricted Boltzmann Machines và thuật toán lan truyền ngược.

➤ **Chương 3: Kết quả thực nghiệm và đánh giá**

Trình bày các kết quả thực nghiệm của hai mô hình SVM và mô hình thống kê học sâu, đưa ra kết quả đánh giá nhận dạng chữ viết tay hạn chế giữa mô hình SVM và mô hình thống kê học sâu.

## Chương 1

### GIỚI THIỆU ĐỀ TÀI

#### 1.1. Giới thiệu về bài toán nhận dạng

Nhận dạng chữ in: đã được giải quyết gần như trọn vẹn (sản phẩm FineReader 11 của hãng ABBYY có thể nhận dạng chữ in theo 192 ngôn ngữ khác nhau, phần mềm nhận dạng chữ Việt in VnDOCR 4.0 của Viện Công nghệ Thông tin – Viện Hàn lâm Khoa học và Công nghệ Việt Nam có thể nhận dạng được các tài liệu chứa hình ảnh, bảng và văn bản với độ chính xác trên 98%).

Nhận dạng chữ viết tay vẫn còn là vấn đề thách thức lớn đối với các nhà nghiên cứu. Bài toàn này chưa thể giải quyết trọn vẹn được vì nó hoàn toàn phụ thuộc vào người viết và sự biến đổi quá đa dạng trong cách viết và trạng thái sức khỏe, tinh thần của từng người viết.

##### 1.1.1. Các giai đoạn phát triển

###### ❖ Giai đoạn 1 (1900 - 1980)

- Nhận dạng chữ được biết đến từ năm 1900, khi nhà khoa học người Nga **Alan Turing** (1912-1954) phát triển một phương tiện trợ giúp cho những người mù.
- Các sản phẩm nhận dạng chữ thương mại có từ những năm 1950, khi máy tính lần đầu tiên được giới thiệu tính năng mới về nhập và lưu trữ dữ liệu hai chiều bằng cây bút viết trên một tấm bảng cảm ứng. Công nghệ mới này cho phép các nhà nghiên cứu làm việc trên các bài toán nhận dạng chữ viết tay online.
- Mô hình nhận dạng chữ viết được đề xuất từ năm 1951 do phát minh của M. Sheppard được gọi là GISMO, một robot đọc-viết.
- Năm 1954, máy nhận dạng chữ đầu tiên đã được phát triển bởi J. Rainbow dùng để đọc chữ in hoa nhưng rất chậm.
- Năm 1967, Công ty IBM đã thương mại hóa hệ thống nhận dạng chữ.

###### ❖ Giai đoạn 2 (1980 - 1990)

- Với sự phát triển của các thiết bị phần cứng máy tính và các thiết bị thu nhận dữ liệu, các phương pháp luận nhận dạng đã được phát triển trong giai đoạn trước