

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



CAO XUÂN TRƯỜNG

**NGHIÊN CỨU MẠNG NƠ-RON NHÂN TẠO
ỨNG DỤNG NHẬN DẠNG CHỮ VIẾT TAY ONLINE**

CHUYÊN NGÀNH: KHOA HỌC MÁY TÍNH

MÃ SỐ: 60.48.01.01

TÓM TẮT LUẬN VĂN THẠC SĨ KỸ THUẬT

HÀ NỘI - 2019

Luận văn được hoàn thành tại:
HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Người hướng dẫn khoa học: GS. TS Từ Minh Phương.
(Ghi rõ học hàm, học vị)

Phản biện 1: PGS. TS. Nguyễn Mạnh Hùng

Phản biện 2: PGS. TS. Nguyễn Hà Nam.

Luận văn được bảo vệ trước Hội đồng chấm luận văn thạc sĩ tại Học viện Công nghệ Bưu chính Viễn thông

Vào lúc: 10 giờ 00 ngày 11 tháng 01 năm 2020

Có thể tìm hiểu luận văn tại:

- Thư viện của Học viện Công nghệ Bưu chính Viễn thông

MỞ ĐẦU

Ngành công nghệ thông tin là một ngành khoa học đang trên đà phát triển mạnh và ứng dụng rộng rãi trên nhiều lĩnh vực. Trong đó nhận dạng chữ viết là một lĩnh vực đem đến nhiều lợi ích thiết thực cho con người. Bài toán nhận dạng chữ viết không chỉ dừng lại ở việc nhận dạng chữ trên giấy mà đã phát triển và mở rộng thành bài toán nhận dạng chữ viết từ các thiết bị điện tử hay còn gọi là nhận dạng chữ viết tay online. Con người từ đó đã có thể viết trực tiếp trên các thiết bị điện tử như smartphone, máy tính bảng hay máy tính xách tay mà không còn phải viết trên giấy như trước đây.

Hiện nay, có nhiều phương pháp khác nhau được áp dụng vào các bài toán nhận dạng. Trong đó việc sử dụng mạng nơ-ron nhân tạo là một trong những phương pháp nhận dạng hiệu quả, có độ chính xác cao. Mạng nơ-ron nhân tạo ra đời xuất phát từ ý tưởng mô phỏng hoạt động của bộ não con người. Kể từ khi ra đời, mạng nơ-ron nhân tạo được ứng dụng rộng rãi trong giải quyết các bài toán về phân loại, dự đoán và nhận dạng bởi khả năng học của nó.

Từ những đặc điểm trên, tôi đã chọn đề tài ***“Nghiên cứu mạng nơ-ron nhân tạo và ứng dụng nhận dạng chữ viết tay online”***. Luận văn này tập trung tìm hiểu phương pháp sử dụng mạng nơ-ron nhân tạo để huấn luyện và nhận dạng chữ viết tay online.

Chương 1. TỔNG QUAN VỀ BÀI TOÁN NHẬN DẠNG CHỮ VIẾT TAY ONLINE

1.1. Giới thiệu bài toán nhận dạng chữ viết tay online

1.1.1. Nhận dạng chữ viết tay online và nhận dạng chữ viết tay offline

Bài toán nhận dạng chữ viết tay offline được đặt ra để nhận dạng các văn bản viết tay trên giấy bằng bút. Với đặc trưng dữ liệu đầu vào là hình ảnh văn bản viết tay được quét hoặc chụp lại. Sau đó, các thuật toán nhận dạng chữ viết tay sẽ được xây dựng dựa trên các hình ảnh này. Các ứng dụng hiện nay về nhận dạng chữ viết tay offline thường quan tâm tới độ chính xác hơn là việc tối ưu thời gian.

Bài toán nhận dạng chữ viết tay online là bài toán nhận dạng chữ viết tay thu được từ các thiết bị điện tử. Với đặc trưng dữ liệu đầu vào là dãy các điểm thu nhận được trong quá trình con người thực hiện viết trên bề mặt các thiết bị điện tử. Yêu cầu là cần có các thiết bị chuyên dụng như màn hình cảm ứng hay các bảng điện tử để ghi lại quá trình di chuyển của nét bút bao gồm các yếu tố như điểm bắt đầu, điểm kết thúc, các điểm trên mặt phẳng mà nét bút đi qua.

1.1.2. Các thuộc tính chữ viết tay online

Chữ viết tay online thu được từ các thiết bị điện tử có một số thuộc tính đặc trưng sau [14]:

Pen-up/ pen-down: pen-up là thuộc tính chỉ nét bút nhấc lên khỏi màn hình và kết thúc một nét, còn pen-down là trạng thái bút chạm màn hình và bắt đầu một nét viết

Speed: Tốc độ viết

X-coordinate: Tọa độ x

Y-coordinate : Tọa độ y

Writing direction: Hướng viết

1.1.3. Phương pháp thu thập chữ viết tay online

**Phương tiện thu thập:*

Việc thu thập dữ liệu chữ viết tay online phải cần các thiết bị chuyên dụng có khả năng ghi nhận các trạng thái của bút như pen up, pen down và tập hợp các điểm mà bút di chuyển qua.

Các thiết bị khác như PDA thì sử dụng các màn hình cảm ứng để ghi nhận sự di chuyển của bút hay tay người dùng.

Khi viết, việc nhận ghi dữ liệu có thể được thu thập theo hai cách khác nhau. Với cách thứ nhất, quỹ đạo bút được chuyển đổi thành hình ảnh điểm ảnh và xử lý với nhận diện ký tự in (OCR) nhận diện như trong nhận dạng off-line. Trong cách thứ hai, quỹ đạo của bút được sử dụng như trình tự của các điểm tọa độ (x, y).

1.1.4. Một số vấn đề khi nhận dạng chữ viết tay online

Những khó khăn trong việc nhận dạng do một số yếu tố sau [1]:

- Hạn chế về số lượng người viết.
- Hạn chế về cách viết.
- Khó khăn về ngôn ngữ.
- Tồn tại nhiều cặp ký tự có hình dạng tương tự, gây khó khăn để nhận biết.
- Tồn tại các nét trề.

1.2. Phương pháp nhận dạng chữ viết tay online

- *Over-segment and classify*: Phương pháp phân đoạn và phân loại.
- *Time-sequence interpretation*: Sự biểu diễn trình tự thời gian.

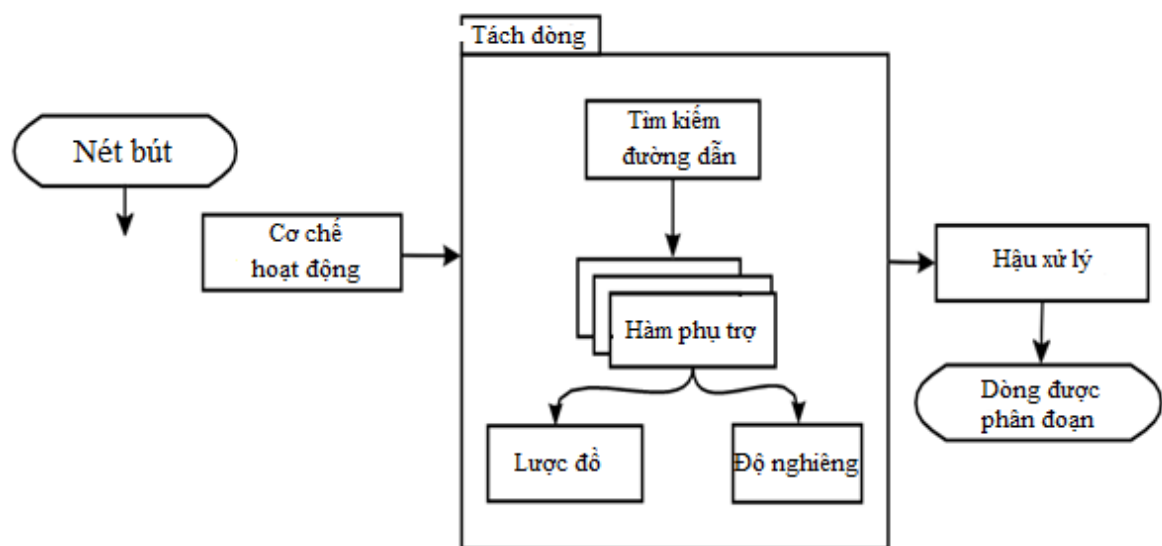
1.3. Quy trình phân đoạn chữ viết tay online

Trong các hệ thống nhận dạng chữ viết tay hoàn chỉnh, yêu cầu được đặt ra là nhận dạng được một trang văn bản hoàn chỉnh. Tuy vậy, trên thực tế, yêu cầu này vẫn đang là một thách thức bởi chưa có một thuật toán nào tối ưu để có thể nhận dạng cả một trang văn bản làm đầu vào. Trước thực tế như vậy, đòi hỏi quá trình nhận dạng phải trải qua các bước phân đoạn. Cụ thể là phân tách các dòng với nhau trong cùng một trang văn bản, tách các từ trên cùng một dòng và phân tích các ký tự riêng biệt của từng từ.

Quá trình tách dòng:

Phân đoạn dòng là bước tiền đề cần thiết trong gần như hầu hết các hệ thống nhận dạng cho văn bản viết tay.

Do dữ liệu online có thể chuyển đổi dễ dàng sang định dạng offline, mọi phương thức nhận dạng offline đều có thể được sử dụng để nhận dạng dữ liệu online. Tuy nhiên để có thể phân đoạn được dữ liệu online nên sử dụng thêm các thông tin online thu thập được. Cách tiếp cận hiệu quả là dựa trên việc tìm ra đường đi tối ưu giữa hai dòng văn bản liên tiếp. Đường dẫn tối ưu giữa hai dòng được tìm thấy bằng cách sử dụng lập trình động. Dữ liệu đầu vào cho thuật toán được đề xuất bao gồm dữ liệu online, nhưng tính đến thông tin offline bổ sung. Hầu hết các bước được thực hiện trên phiên bản offline của dữ liệu.



Hình 1.3. Quy trình tách dòng

Tiền xử lý:

Ở bước tách dòng, tiền xử lý các dữ liệu thu được là bước quan trọng trong hệ thống nhận dạng chữ viết tay bởi mỗi người viết có các phong cách viết khác nhau về độ nghiêng, độ xiên, chiều cao, chiều rộng của các ký tự. Phương pháp thường được dùng trong bước tiền xử lý ở một số hệ thống nhận dạng như sau:

- Tiền xử lý trực tiếp:

- Chuẩn hóa qua các bước:
- + Các phần được hiệu chỉnh đưa về đúng độ nghiêng của chúng.
- + Bước chuẩn hóa về độ xiên được thực hiện dựa vào việc sử dụng lược đồ xám.
- + Bước tiếp theo là tính đường cơ sở và đường corpus.
- + Bước cuối cùng của việc chuẩn hóa là chuẩn hóa độ rộng ký tự.

1.4. Hướng tiếp cận bài toán nhận dạng chữ viết tay online

Đối với bài toán nhận dạng chữ viết tay online có thể tiếp cận theo hai hướng:

Cách tiếp cận thứ nhất: Sử dụng các mô hình học máy để huấn luyện một bộ phân lớp cho toàn bộ các từ có thể từ điển. Tuy nhiên cách tiếp cận này chỉ có thể áp dụng với từ điển có số lượng từ nhỏ (khoảng 100-200). Đối với những từ điển có số lượng từ lớn hơn thì ngay cả việc thu thập dữ liệu cũng gặp khó khăn vì sự đa dạng trong cách thể hiện của mỗi từ.

Cách tiếp cận thứ hai: Tiến hành phân đoạn cả từ rồi sử dụng một bộ phân lớp ở mức ký tự để tìm ra kết quả cuối cùng. Ở cách tiếp cận này chỉ có thể áp dụng nếu đặt ra những ràng buộc nhất định như có khoảng cách rõ ràng giữa các ký tự trong cùng một từ, v.v.

1.5. Ứng dụng bài toán nhận dạng chữ viết tay online

Trong bối cảnh các thiết bị cảm ứng như Smartphone, Ipad, v.v ngày càng trở nên phổ biến, thì bài toán nhận dạng chữ viết online trở nên đặc biệt được quan tâm. Con người có thể viết trực tiếp trên một màn hình tinh thể lỏng (LCD) hiển thị với một cây bút điện tử hay đơn giản là ngón tay. Màn hình là một ma trận nhạy cảm ghi lại sự chuyển động của đầu bút trên bề mặt. Quỹ đạo của các cây bút xuất hiện gần như ngay lập tức trên màn hình. Từ đó, bài toán nhận dạng chữ viết tay online đóng vai trò quan trọng, cho phép máy tính có thể hiểu người viết muốn thực hiện thao tác gì, lệnh gì, v.v.

Ngày nay có thể kể tới một số ứng dụng cụ thể của nhận dạng chữ viết tay online như:

- Hệ thống nhận dạng chữ viết tay rời rạc online trên một trạm làm việc của IBM do H.S.M.Beigi, C.C.Tapert, M.Ukeison và C.G.Wolf ở phòng thực hành Watson IBM cài đặt.

- Google đã tích hợp vào dịch vụ Google Translate.

- Bên cạnh đó, Google còn tung ra một sản phẩm nhận dạng chữ viết tay khá chính xác là ứng dụng Handwritng Input.

1.6. Kết luận chương 1

Chương 1 đã trình bày bài toán nhận dạng chữ viết tay, các đặc trưng của bài toán nhận dạng chữ viết tay online, các thuộc tính cơ bản của chữ viết online, một số phương pháp thu thập và một số vấn đề khó khăn gặp phải khi nhận dạng chữ viết tay online.

Bài toán nhận dạng chữ viết tay online phải trải qua các bước phân đoạn như phân đoạn dòng, từ và ký tự, nhưng trong luận văn chỉ tập trung nghiên cứu về nhận dạng ký tự. Ngoài ra, nội dung chương cũng trình bày các ứng dụng phổ biến của bài toán nhận dạng chữ viết tay online hiện nay.

Chương 1 cũng trình bày tổng quan về một số phương pháp nhận dạng chữ viết online đang được quan tâm và nghiên cứu hiện nay, trong đó mô hình mạng nơ ron nhân tạo kết hợp với các kỹ thuật học sâu là đối tượng nghiên cứu ở chương 2 và là công cụ nhận dạng ở chương 3 của luận văn.

Chương 2. ỨNG DỤNG MẠNG NƠ RON NHÂN TẠO VÀO BÀI TOÁN NHẬN DẠNG CHỮ VIẾT TAY ONLINE

2.1. Mạng nơ ron nhân tạo

2.1.1. Các khái niệm

Mạng nơ ron nhân tạo (Artificial Neural Network) gọi tắt là mạng nơ ron (neural network), là mô hình xử lý thông tin phỏng theo cách thức xử lý thông tin của các hệ nơ ron sinh học trong bộ não người. Nó được tạo nên từ một số lượng lớn các phân tử (gọi là phân tử xử lý hay nơ ron) kết nối với nhau thông qua các liên kết (gọi là trọng số liên kết) làm việc như một thể thống nhất để giải quyết một vấn đề cụ thể nào đó.

2.1.2. Đặc trưng của mạng nơ ron

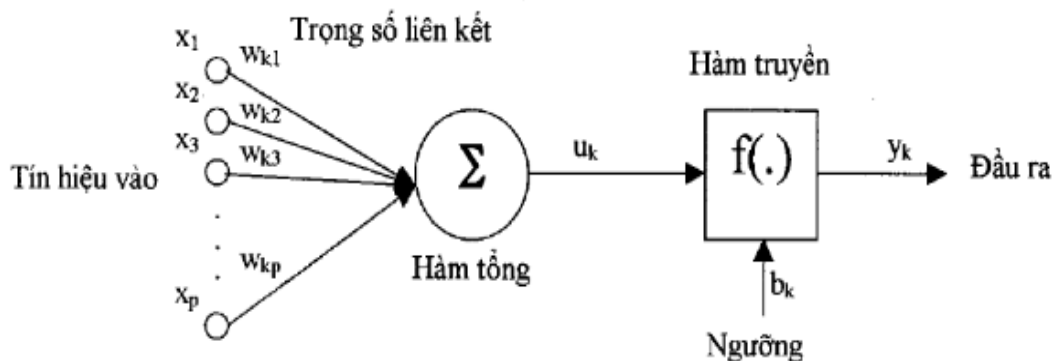
Một số đặc trưng của mạng nơ ron thu được từ việc mô phỏng trực tiếp bộ não con người như sau:

- *Tính chất phi tuyến:*
- *Tính chất thích nghi:*
- *Tính chất đưa ra lời giải có bằng chứng:*
- *Tính chất chấp nhận sai sót:*
- *Tính chất đồng dạng trong phân tích và thiết kế.*

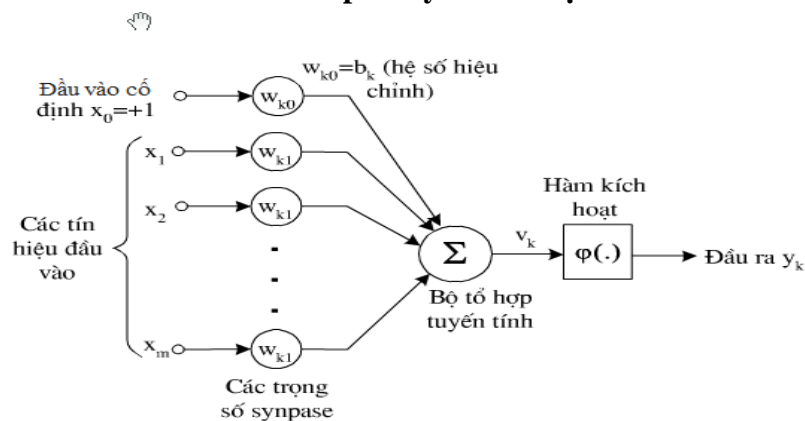
2.1.3. Kiến trúc mạng nơ ron nhân tạo

Từ định nghĩa cho thấy đơn vị hình thành nên mạng nơ ron là các nơ ron. Chức năng của mạng nơ ron nhân tạo tương tự như chức năng đơn giản nhất của các nơ ron sinh học đó là tích lũy (ghi nhớ) những tri thức đã học về các sự vật để nhận biết, phân biệt chúng mỗi khi gặp lại. Chức năng này được thực hiện thông qua một chuỗi liên tiếp các quá trình xử lý thông tin của các nơ ron trong mạng.

. Sơ đồ 2.1 chỉ ra mô hình cấu trúc một nơ ron nhân tạo:



Hình 2.1. Mô hình phi tuyến của một nơ ron

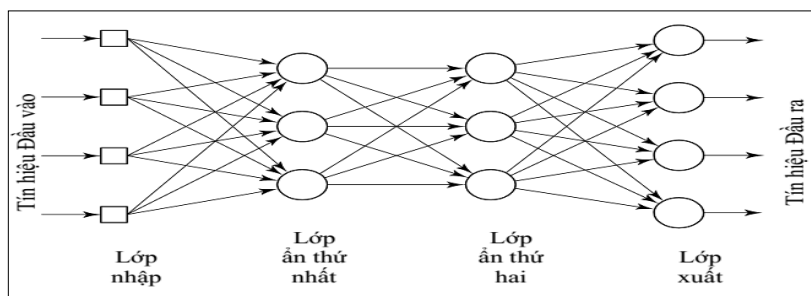


Hình 2.2. Mô hình phi tuyến thứ hai của một nơ ron

2.1.4. Phân loại mạng nơ ron nhân tạo

2.1.4.1. Mạng truyền thẳng

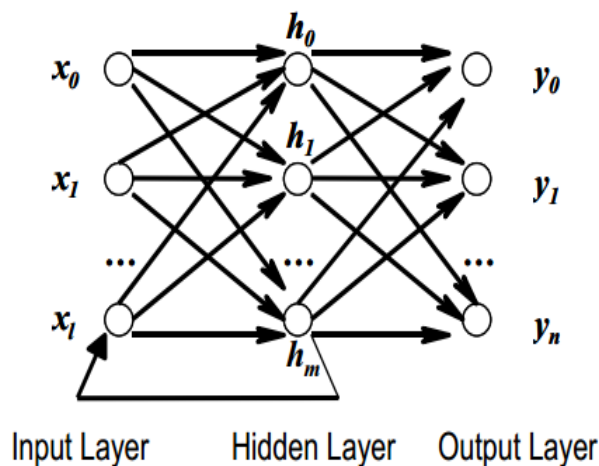
Dòng dữ liệu từ đơn vị đầu vào đến đơn vị đầu ra chỉ được truyền thẳng. Việc xử lý dữ liệu có thể mở rộng ra nhiều lớp, nhưng không có các liên kết phản hồi.



Hình 2.3. Mạng nơ ron truyền thẳng nhiều lớp

2.1.4.2. Mạng hồi quy

Mạng hồi quy là kiểu kiến trúc mạng không chỉ truyền dữ liệu theo một chiều, mà còn truyền dữ liệu từ bước sau ngược trở lại bước trước.



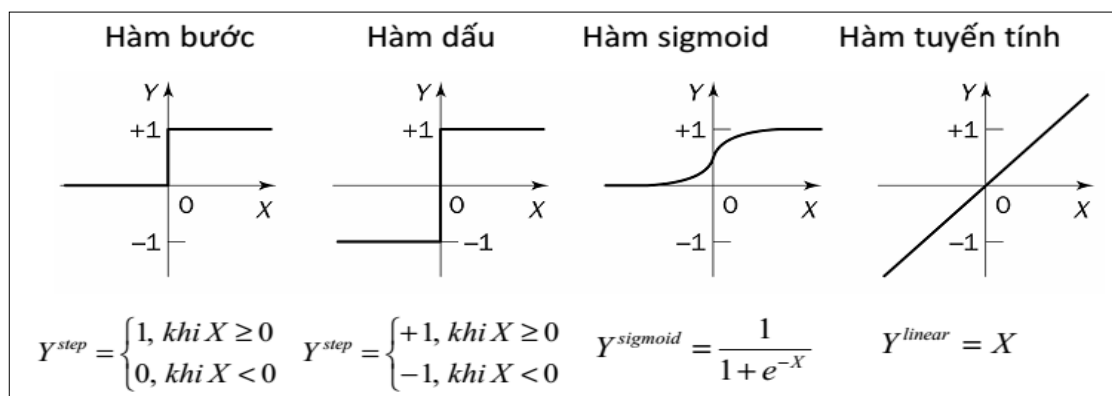
Hình 2.4. Mô hình mạng hồi quy

2.1.5. Hàm kích hoạt

Hàm kích hoạt, ký hiệu bởi $\varphi(v)$, được dùng để xác định đầu ra của mạng nơ ron. Các hàm kích hoạt phải có các đặc điểm sau:

- + Hàm bị chặn trên và chặn dưới
- + Hàm có tính đơn điệu
- + Hàm phải có tính liên tục

Dưới đây là một số kiểu hàm kích hoạt cơ bản thường được sử dụng:



Hình 2.5. Các hàm kích hoạt tiêu biểu

2.1.6. Phương pháp huấn luyện mạng nơ ron

2.1.6.1. Học có giám sát

Với phương pháp này, cần đưa vào bộ dữ liệu huấn luyện và kết quả đầu ra mong muốn tương ứng. Thuật toán sẽ tiến hành tính toán trọng số để mỗi vector đầu vào sẽ có kết quả đầu ra đúng với kết quả đầu ra mong muốn. Quá trình học có giám sát cụ thể như sau:

2.1.6.2. Học không có giám sát

Học không có giám sát là học với tập dữ liệu huấn luyện ban đầu hoàn toàn chưa được gán nhãn. Đây là phương pháp học sử dụng cho lớp bài toán phân cụm. Quá trình huấn luyện của phương pháp học không có giám sát có đặc điểm như sau:

- + Sử dụng bộ dữ liệu huấn luyện chưa được gán nhãn (không có thông tin về giá trị đầu ra mong muốn).
- + Lựa chọn thuật toán học không có giám sát để huấn luyện.
- + Sử dụng phương pháp thử nghiệm để đánh giá chất lượng.

2.1.6.3. Học tăng cường

Trong luật học có giám sát, các giá trị đầu ra được biết chính xác đối với mỗi đầu vào. Tuy nhiên, trong thực tế có một số trường hợp chỉ biết ít thông tin chi tiết, chẳng hạn mạng chỉ biết rằng giá trị đầu ra thực sự quá cao hay có thể mạng chỉ có được thông tin phản hồi báo rằng đầu ra đúng hay sai. Thuật học dựa trên thông tin đánh giá này được gọi là thuật học củng cố hay học tăng cường, thông tin phản hồi được gọi là tín hiệu tăng cường.

2.1.7. Mạng nơ ron truyền thẳng một lớp ẩn và thuật toán lan truyền ngược

2.1.7.1. Mạng nơ ron truyền thẳng một lớp ẩn

Mạng nơ ron truyền thẳng một lớp ẩn là mạng nơ ron truyền thẳng gồm 3 lớp, trong đó có một lớp đầu vào, một lớp ẩn và một lớp đầu ra. Các nơ ron ở mỗi lớp liên kết với nơ ron ở lớp kế tiếp theo một chiều từ lớp đầu vào đến lớp ẩn và đến nơ ron lớp đầu ra, không có liên kết ngược lại hoặc liên kết giữa các nơ ron trong một lớp.

2.1.7.2. Thuật toán lan truyền ngược

Phương pháp huấn luyện phổ biến cho mạng nơ ron truyền thẳng một lớp ẩn là thuật toán lan truyền ngược. Thuật toán lan truyền ngược là dạng tổng quát của thuật toán trung bình phương tối thiểu (Least Means Square –LMS).

Huấn luyện mạng nơ ron sử dụng thuật toán Lan truyền ngược gồm hai quá trình: Quá trình truyền tuyến tính và quá trình truyền ngược:

2.2. Ứng dụng mạng nơ ron vào bài toán nhận dạng chữ viết tay online

2.2.1. Thu thập dữ liệu

Chương trình được thực nghiệm trên bộ dữ liệu UNIPEN.

**Một số đặc điểm về cơ sở dữ liệu UNIPEN [16]:*

Định dạng UNIPEN là định dạng ASCII được thiết kế đặc biệt cho dữ liệu được thu thập với bất kỳ thiết bị cảm ứng nhạy cảm, điện trở hay điện từ nào cung cấp thông tin quỹ đạo được viết bằng bút điện từ. Người dùng có thể chuyển đổi định dạng của riêng họ sang hoặc từ định dạng UNIPEN hoặc thu thập dữ liệu trực tiếp từ định dạng đó.

Chương trình thực nghiệm được xây dựng với bộ cơ sở dữ liệu có 50000 mẫu được sử dụng, trong đó có 70% dùng cho bộ Train và 30% cho bộ Test, bao gồm nhiều mẫu viết của nhiều người khác nhau trên thế giới được lựa chọn ngẫu nhiên để đưa vào huấn luyện.

2.2.2. Mô hình của bài toán nhận dạng chữ viết tay online sử dụng mạng nơ ron nhân tạo

Mô hình tổng quát của hệ thống nhận dạng chữ viết tay online sử dụng mạng nơ ron nhân tạo được chia thành hai giai đoạn: Giai đoạn huấn luyện và giai đoạn nhận dạng.

- Giai đoạn huấn luyện gồm các bước chính: Lựa chọn cơ sở dữ liệu huấn luyện. Với từng file cơ sở dữ liệu tiến hành trích chọn đặc trưng rồi đưa vào mạng nơ ron để huấn luyện.

- Giai đoạn nhận dạng: vẽ/viết ký tự lên màn hình thiết bị nhận dạng, sau đó tiến hành trích chọn đặc trưng. Những đặc trưng này sẽ được đưa vào mạng nơ ron đã huấn luyện ở trên để tiến hành nhận dạng.

2.2.3. Trích chọn đặc trưng và xây dựng vector đầu vào

Trích chọn đặc trưng là quá trình tìm ra các thông tin hữu ích và đặc trưng nhất cho mẫu đầu vào để sử dụng cho quá trình huấn luyện và nhận dạng.

Với luận văn, việc trích chọn đặc trưng các mẫu đưa vào được thực hiện dựa trên việc đếm số điểm đen trên mặt phẳng nơi nét bút đi qua, từ đó mạng nơ ron lưu lại các đặc trưng cơ bản này cho những lần trích chọn đặc trưng khác của các mẫu khác với cùng một ký tự được nhận dạng và của các ký tự khác nhau

Bảng 2.1. Kết quả thực nghiệm với vector đầu vào khác nhau

m	n	Số phần tử của vector đầu vào ($m*n$ hoặc $m*n+1$)	Kết quả nhận dạng
4	2	8	40%
4	2	9	45%
5	2	10	58%
5	2	11	82%
3	4	13	76%
3	4	12	73%
7	2	15	64%
7	2	14	62%
4	4	17	54%
4	4	16	53%
8	2	17	49%
8	2	16	45%

2.2.4. Các tham số của mạng nơ ron.

2.2.4.1. Số phần tử ở các lớp

- Lớp đầu vào: Số lượng node được lựa chọn là 11.

- Lớp ẩn: Số lượng node ở tầng ẩn được lựa chọn ngẫu nhiên thỏa mãn điều kiện:

- Lớp đầu ra: Số lượng node của lớp đầu ra tương ứng với số lượng ký tự khác nhau được đưa vào huấn luyện.

2.2.4.2. Hệ số học (Learning Rate)

Learning rate hay hệ số học là một trong những thông số mà điều chỉnh để một mạng nơ ron học nhanh và huấn luyện hiệu quả.

2.2.4.3. Bước đà (Momentum)

Việc sử dụng Momentum (bước đà) có tác dụng giảm các hậu quả hạn chế từ việc sử dụng hệ số học không như muốn.

Về bản chất, trong phần đầu của quá trình huấn luyện, bước đà chỉ có tác dụng giúp quá trình trở nên nhanh hơn. Sau đó khi quá trình huấn luyện gần đạt tới hội tụ.

Nhược điểm của bước đà là giá trị của nó là ngẫu nhiên và phải được sử dụng thử và sai để tìm ra giá trị tốt.

2.2.5. Huấn luyện mạng nơ ron

Mạng nơ ron được sử dụng trong báo cáo luận văn là mạng nơ ron truyền thẳng một lớp ẩn. Trong đó lớp đầu vào và lớp ẩn có số lượng nodes cho phép tùy chỉnh được, để sau quá trình kiểm tra độ chính xác của mạng nơ ron có thể xác định được cấu hình mạng tốt nhất.

Ở bước huấn luyện, mạng nơ ron sử dụng thuật toán lan truyền ngược backpropagation.

2.2.6. Nhận dạng chữ viết tay online sử dụng mạng nơ ron

Trong giai đoạn nhận dạng, dữ liệu chữ viết tay thu được sẽ được trích chọn đặc trưng. Những đặc trưng này sau đó được đưa vào mạng nơ ron đã được trải qua quá trình huấn luyện để tiến hành nhận dạng chữ.. Kết quả của bước nhận dạng ký tự sẽ được sử dụng để đánh giá hiệu năng của mạng nơ ron trong nhận dạng chữ viết tay online.

2.3. Kết luận chương 2

Chương 2 trình bày khái quát về mạng nơ ron nhân tạo, về đặc trưng, kiến trúc, các phương pháp huấn luyện mạng và đặc biệt tìm hiểu về mạng nơ ron truyền thẳng một lớp ẩn và thuật toán lan truyền ngược.

Tiếp đó, trình bày 02 mô hình nhận dạng trong quá trình huấn luyện và quá trình nhận dạng chữ. Sau đó là trình bày về quá trình trích chọn các đặc trưng của chữ viết và các tham số đặc trưng trong quá trình huấn luyện và nhận dạng. Quá trình trích chọn đặc trưng được khai thác trên các tọa độ (x,y) trên mặt phẳng mà nét bút đi qua, đồng thời có thêm thông tin của số nét bút.

Luận văn đã xây dựng và tiến hành thực nghiệm và cho kết quả về số lượng tham số đầu vào hợp lý nhất. Cuối cùng, đưa ra cái nhìn khái quát về quá trình huấn luyện và nhận dạng chữ viết online sử dụng mạng nơ ron nhân tạo. Nội dung các chương 1 và chương 2 là tiền đề cho việc xây dựng chương trình thực nghiệm nhận dạng chữ viết tay online ở mức độ nhận dạng ký tự được trình bày ở chương 3 của luận văn.

Chương 3. CÀI ĐẶT THỰC NGHIỆM VÀ ĐÁNH GIÁ KẾT QUẢ

3.1. Cài đặt thực nghiệm

3.1.1. Ngôn ngữ lập trình

Chương trình thực nghiệm nhận dạng chữ viết tay online sử dụng mạng nơ ron được lập trình bằng ngôn ngữ lập trình C# trên Visual Studio 2013.

3.1.2. Cấu hình máy

Bộ vi xử lý: Intel(R) Core(TM) i3-3120M; CPU: 2.50GHz; RAM: 6GB

Hệ điều hành: Windows 7 Ultimate.

3.1.3. Cấu hình mạng nơ ron

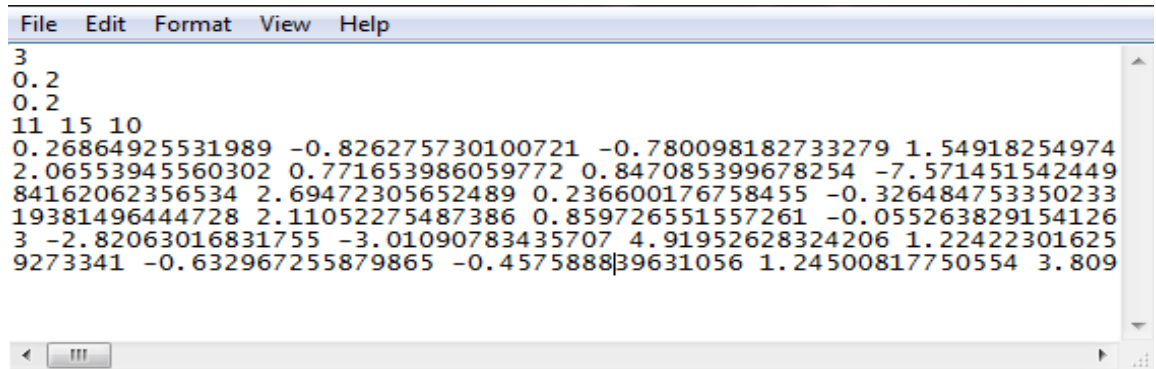
- Số node input: 11.
- Các tham số còn lại gồm: Số node tầng ẩn, số lượng vòng lặp, ngưỡng MSE, momentum (alpha) và learning rate (beta) đều cho phép tùy chỉnh thay đổi để tìm ra cấu hình mạng tốt nhất.
- Alpha (momentum): Nằm trong phạm vi $0 \leq \alpha < 1$.
- Beta (learning rate): Nằm trong phạm vi $0 \leq \beta < 1$.

3.1.4. Giới thiệu chương trình

- Huấn luyện mạng nơ ron:

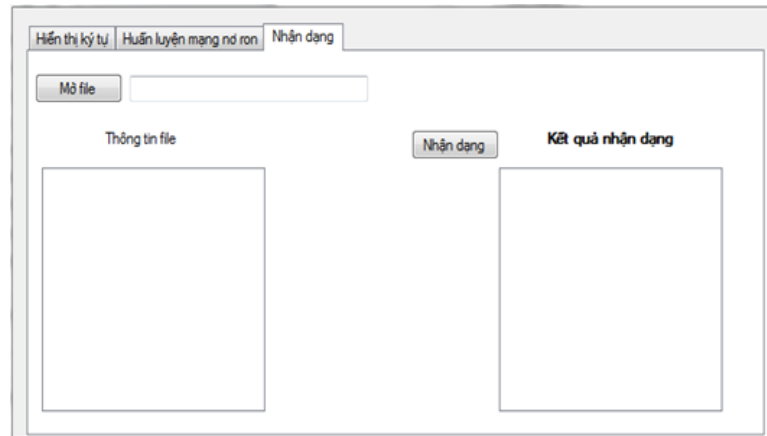
Hình 3.1. Giao diện huấn luyện mạng nơ ron

- Lưu mạng đã huấn luyện:



Hình 3.2. Các giá trị của mạng được lưu lại sau huấn luyện

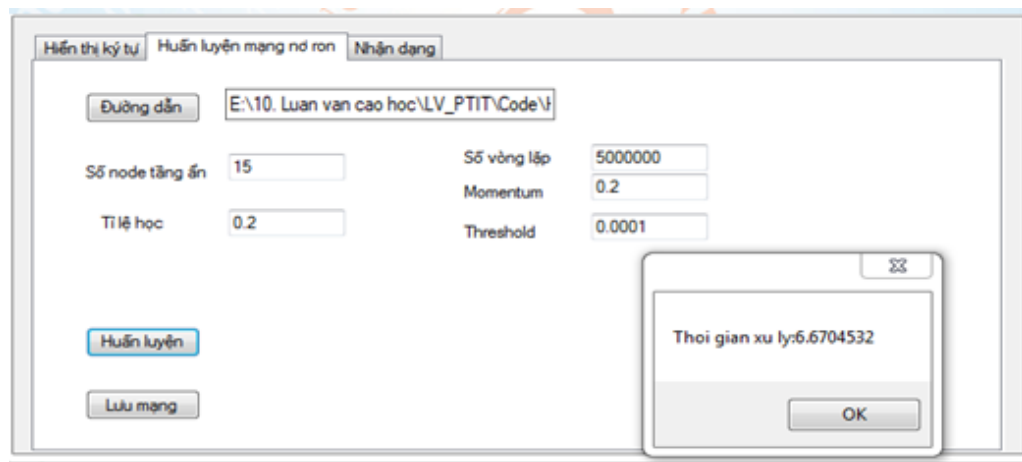
- Nhận dạng chữ viết trong cơ sở dữ liệu UNIPEN:



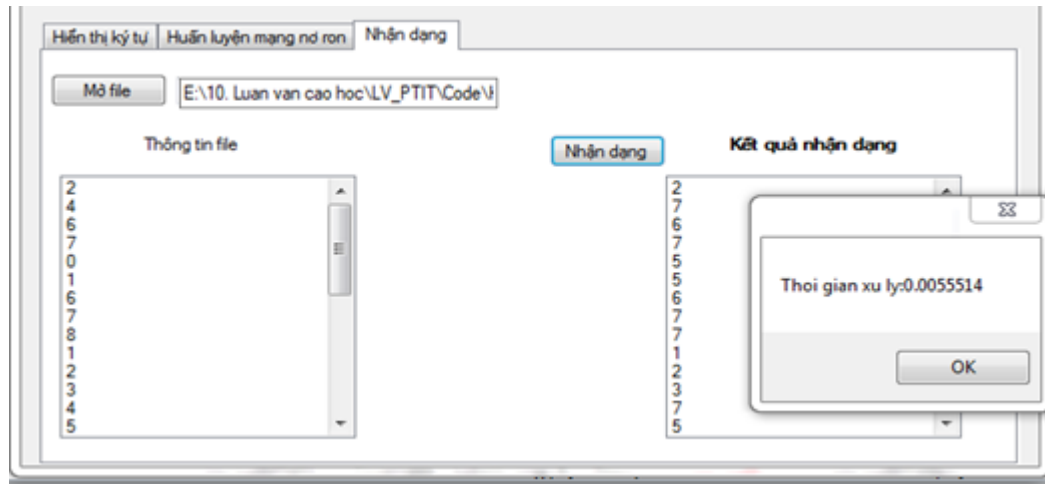
Hình 3.3. Giao diện nhận dạng chữ viết online

3.2. Đánh giá kết quả thực nghiệm

3.2.1. Kết quả thực nghiệm



Hình 3.4. Demo thời gian huấn luyện



Hình 3.5. Demo minh họa thời gian nhận dạng

3.2.2. Đánh giá kết quả

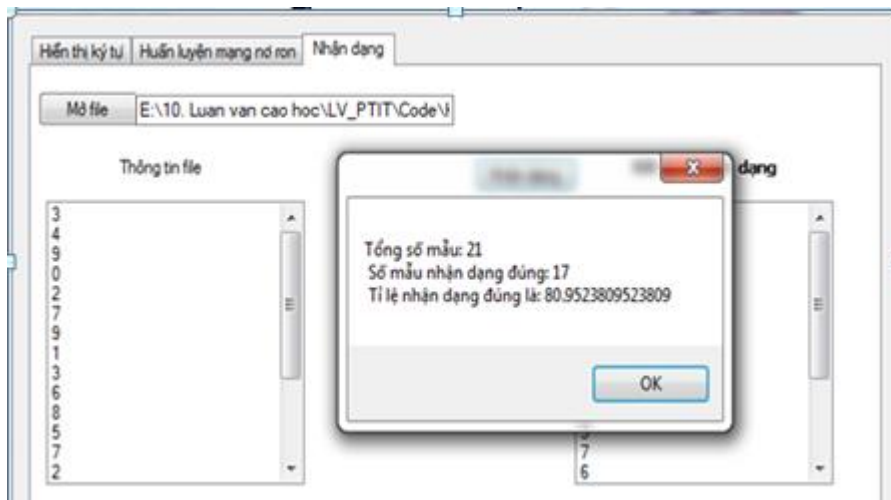
Kết quả nhận dạng cho từng ký tự:

Bảng 3.1. Thống kê kết quả nhận dạng ký tự

Ký tự	Số lượng mẫu để kiểm tra	Kết quả nhận dạng đúng
a	547	0.84
b	269	0.81
c	250	0.95
d	672	0.88
e	582	0.93
f	182	0.82
g	240	0.92
h	250	0.83
i	578	0.73
j	283	0.74
k	238	0.78
l	243	0.64
m	248	0.91
n	237	0.87
o	547	0.92

p	234	0.87
q	256	0.82
r	243	0.37
s	239	0.66
t	245	0.72
u	579	0.66
v	262	0.88
w	178	0.90
x	256	0.84
y	218	0.82
z	154	0.74
0	231	0.76
1	167	0.78
2	342	0.67
3	332	0.68
4	367	0.79
5	453	0.87
6	342	0.85
7	321	0.78
8	423	0.81
9	123	0.77
Tỷ lệ trung bình		0.82

Kết quả nhận dạng khi load file chứa nhiều ký tự khác nhau:



Hình 3.6. Kết quả nhận dạng khi load file nhiều ký tự

3.2.3. Đánh giá kết quả

Sau khi thu được các kết quả thực nghiệm, cho thấy tỷ lệ nhận dạng trung bình đạt 80%. Trong đó, phương pháp nhận dạng chữ viết tay sử dụng mạng nơ ron nhân tạo với thuật toán lan truyền ngược, đạt được kết quả tốt ở các ký tự “c”, “d”, “e”, “o”, “p”, “q”, “l”, “3”, “6”, “5”, “9”. Một số ký tự khác như “l” thường nhận nhầm thành số “1”, “u” thành “v”, “v” thành “u” hay các số “4” và “1” cũng hay bị nhận dạng nhầm với nhau v.v.

3.3. Kết luận chương 3

Nội dung chương 3 trình bày về quá trình nhận dạng chữ viết online ở mức độ nhận dạng ký tự được xây dựng trong chương trình mô phỏng. Bằng các kiến thức được nghiên cứu và tìm hiểu ở chương 1 và chương 2, chương 3 đã xây dựng thành công phần mềm nhận dạng chữ viết tay online ở cấp nhận dạng ký tự sử dụng mạng nơ ron nhân tạo trên bộ cơ sở dữ liệu UNIPEN.

KẾT LUẬN

1. Các kết quả nghiên cứu của luận văn

Về mặt lý thuyết, luận văn đã trình bày được một số nội dung cơ bản sau:

- Giới thiệu tổng quan về bài toán nhận dạng chữ viết tay, trong đó trọng tâm là bài toán nhận dạng chữ viết tay online, những đặc trưng cơ bản của chữ viết online và các phương pháp thu thập chữ viết tay online.

- Giới thiệu tổng quan về mạng nơ ron, kiến trúc và các phương pháp huấn luyện mạng nơ ron. Cụ thể về mạng nơ ron truyền thẳng một lớp ẩn và thuật toán lan truyền ngược.

Về mặt thực nghiệm:

- Xây dựng được một mạng nơ ron truyền thẳng một lớp ẩn sử dụng thuật toán lan truyền ngược ứng dụng nhận dạng chữ viết tay online ở mức độ nhận dạng ký tự sử dụng cơ sở dữ liệu UNIPEN.

2. Hướng phát triển của luận văn

- Tiếp tục cải thiện và kết hợp nhiều thuật toán để có thể trực tiếp nhận dạng được chữ viết online với các mức độ cao hơn (nhận dạng được chữ viết, câu văn, đoạn văn v.v) với độ chính xác cao hơn.

- Nghiên cứu thêm các thuật toán trích chọn đặc trưng để tỷ lệ nhận dạng đúng cao hơn.

- Xây dựng được phần mềm nhận dạng trực tiếp từ quá trình viết lên các thiết bị điện tử.

TÀI LIỆU THAM KHẢO

Tiếng Việt

- [1] Nguyễn Hoàng Hải (2012), *Hệ thống nhận dạng chữ viết tay in hoa trực tuyến*, Luận văn thạc sĩ khoa học, tr. 14 – 15.

Tiếng Anh

- [2] A. Graves, M. Liwicki, S. Fernandez, R. Bertolami, H. Bunke, and J. Schmidhuber (2009), A novel connectionist system for unconstrained handwriting recognition, *IEEE Trans. Pattern Anal. Mach. Intell.*, Tập 31, Số 5, tr. 855-868.
- [3] A. Delaye and E. Anquetil (2013), HBF49 feature set: A first unified baseline for online symbol recognition, in *Proc. Pattern Recognit*, Tập 46, Số 1, tr. 117–130.
- [4] C. C. Tappert, C. Y. Suen, and T. Wakahara (1990), “The state of the art in online handwriting recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, Tập 12, Số 8, tr. 787–808.
- [5] Daniel Keysers, Thomas Deselaers, Henry A. Rowley, Li-Lun Wang, and Victor Carbune, (2017) Multi-Language Online Handwriting Recognition, *IEEE Trans. Pattern Anal. Mach. Intell.*, Tập 39, Số 6, tr. 63–84.
- [6] G. Nagy (2000), Twenty Years of Document Image Analysis in PAMI, *IEEE Trans. Pattern Analysis and Machine Intelligence*, tr. 135-138.
- [7] G. Vamvakas, B. Gatos, I. Pratikakis, N. Stamatopoulos, A. Roniotis and S.J. Perantonis (2007), Hybrid Off-Line OCR for Isolated Handwritten Greek Characters, *The Fourth IASTED International Conference on Signal Processing, Pattern Recognition, and Applications*, tr. 197-202.
- [8] J. A. Pittman (2007), Handwriting recognition: Tablet PC text input, *IEEE Comput.*, Tập 40, Số 9, tr. 49–54.
- [9] J. H. Kim and B. Sin (2014), Online handwriting recognition, in *Proc. Handbook Doc. Image Process. Recognit.*, tr. 887–915.

- [10] J.Makhoul, T. Starner, R. Schwartz, and G. Chou (1994), Online Cursive Handwriting Recognition Using Speech Recognition Methods, *Proc. ICASSP'94*, tr. 125-128.
- [11] J.Schenk and G. Rigoll (2006), Novel hybrid NN/HMM modelling techniques for on-line handwriting recognition, *in Proc. Int. Workshop Frontiers Handwriting Recognit*, tr. 619–623.
- [12] K.-F. Chan, D.-Y. Yeung (1998), Elastic Structural Matching for Online Handwritten Alphanumeric Character Recognition, *Proc. 14th Int. Conf. Pattern Recognition, Brisbane, Australlia*, tr. 1508-1511.
- [13] L.Yaeger, B.Webb, and R.Lyon (1998), Combining nơ ron networks and context-driven search for on-line, printed handwriting recog- nition in the Newton, *in AAAI AI Magazine. Berlin, Germany, Springer*.
- [14] Marcus Liwicki and Horst Bunke (2008), RECOGNITION OF WHITEBOARD NOTES Online, Offline and Combination, tr. 129 -130.
- [15] R. Plamondon and S. N. Srihari (200), On-line and off-line handwriting recognition: A comprehensive survey, *IEEE Trans. Pattern Anal. Mach. Intell.*, Tập 22, Số 1, tr. 63–84.
- [16] S. Jaeger, S. Manke, J. Reichert, and A. Waibel (2001), Online handwrit- ing recognition: The NPen++ recognizer, *Int. J. Doc. Anal. Recog- nit.*, Tập 3, Số 3, tr. 169–180.
- [17] Simon Haykin (2009), Nơ ron networks and learning machines 3rd edition, *Pearson education*, tr. 34 – 37.