

**ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ**

Nông Thị Hoa

**CẢI TIẾN QUÁ TRÌNH HỌC
CỦA MỘT SỐ MẠNG NƠ-RON GHI NHỚ**

Chuyên ngành: Khoa học máy tính

Mã số: 62.48.01.01

LUẬN ÁN TIẾN SĨ KHOA HỌC MÁY TÍNH

NGƯỜI HƯỚNG DẪN KHOA HỌC

PGS.TS. Bùi Thế Duy

Hà Nội – 2015

Lời cam đoan

Tôi xin cam đoan đây là công trình nghiên cứu của riêng tôi. Các kết quả được viết chung với các tác giả khác đều được sự đồng ý của các đồng tác giả trước khi đưa vào luận án. Các kết quả nêu trong luận án là trung thực và chưa từng được ai công bố trong các công trình nào khác.

Tác giả

Lời cảm ơn

Luận án được thực hiện tại Trường Đại học Công nghệ, Đại học Quốc gia Hà Nội, dưới sự hướng dẫn của PGS.TS. Bùi Thế Duy.

Tôi xin bày tỏ lòng biết ơn sâu sắc tới PGS.TS. Bùi Thế Duy và GS.TS. Đặng Quang Á, các thầy đã có những định hướng giúp tôi thành công trong việc nghiên cứu của mình. Các thầy cũng đã động viên và chỉ bảo giúp tôi vượt qua những khó khăn để tôi hoàn thành được luận án này. Tôi cũng chân thành cảm ơn TS.Võ Đình Bảy, TS. Đặng Trung Kiên, Ths. Nguyễn Quốc Đại, những người đã cho tôi nhiều kiến thức quý báu về viết bài báo khoa học và trợ giúp xuất bản các bài báo. Những sự chỉ bảo quý giá của các thầy, đồng nghiệp đã giúp tôi hoàn thành tốt luận án này.

Tôi cũng xin cảm ơn tới các Thầy, Cô thuộc Khoa Công nghệ thông tin, Trường Đại học Công nghệ, Đại học Quốc gia Hà Nội đã tạo mọi điều kiện thuận lợi giúp tôi trong quá trình làm nghiên cứu sinh.

Tôi cũng xin cảm ơn ban lãnh đạo trường Đại học Công nghệ thông tin và truyền thông, Đại học Thái Nguyên đã tạo mọi điều kiện về mặt thời gian và công tác chuyên môn trong quá trình làm nghiên cứu sinh.

Cuối cùng, tôi xin gửi lời cảm ơn sâu sắc tới gia đình và bạn bè, những người đã luôn ủng hộ và hỗ trợ tôi về mọi mặt để tôi yên tâm học tập và đạt được kết quả học tập tốt.

MỤC LỤC

Lời cam đoan	2
Lời cảm ơn	3
MỤC LỤC	4
Danh mục các từ viết tắt	8
Danh mục các bảng	9
Danh mục các hình vẽ, đồ thị	11
Danh mục các thuật toán.....	13
Danh mục các định lý	13
MỞ ĐẦU	14
CHƯƠNG 1. MẠNG NƠ-RON NHÂN TẠO	17
1.1 Nơ-ron sinh học.....	17
1.2 Nơ-ron nhân tạo	18
1.3 Mạng nơ-ron nhân tạo	19
1.4 Các luật học của ANN.....	22
1.5 Ưu và nhược điểm của ANN.....	24
1.6 Ứng dụng của ANN	24
1.7 Kết luận chương	26
CHƯƠNG 2. MỘT SỐ MẠNG NƠ-RON GHI NHỚ	27
2.1 Logic mờ	27
2.1.1 Định nghĩa	27
2.1.2 Các phép toán với tập mờ.....	27
2.2 Toán học hình thái.....	28
2.2.1 Lưới đầy đủ.....	28
2.2.2 Các thao tác cơ bản với lưới đầy đủ	28
2.3 Mô hình AM.....	29
2.3.1 Khái niệm về AM	29
2.3.2 Hoạt động của AM	29
2.3.3 Một số đặc điểm của AM	30
2.4 Mô hình BAM	31

2.4.1 Mạng Hopfield.....	31
2.4.2 Khái niệm về BAM	33
2.4.3 Quá trình học của BAM	34
2.4.4 Quá trình nhớ lại của BAM	35
2.4.5 Hàm năng lượng của BAM.....	35
2.4.6 Chiến lược học nhiều lần dùng số lần lặp tối thiểu để học một cặp mẫu	36
2.5 Mô hình FAM	36
2.5.1 Khái niệm FAM.....	36
2.5.2 Các kiểu nơ-ron trong FAM	37
2.5.3 Các FAM của Kosko và sự tổng quát hóa	38
2.6 Mô hình ART	39
2.6.1 Cấu trúc của ART	39
2.6.2 Các bước hoạt động chính của ART.....	40
2.6.3 Họ các mô hình của ART	41
2.7 Mô hình Fuzzy ART	41
2.7.1 So sánh với ART	41
2.7.2 Thuật toán Fuzzy ART	42
2.7.3 Fuzzy ART với mã hóa đầy đủ.....	43
2.7.3 Thước đo chất lượng phân cụm.....	44
2.8 Kết luận chương	44
CHƯƠNG 3. THUẬT TOÁN HỌC CẢI TIẾN CHO BỘ NHỚ LIÊN KẾT HAI CHIỀU	45
3.1 Giới thiệu chung.....	45
3.2 Các nghiên cứu liên quan	45
3.2.1 Các mô hình lý thuyết.....	45
3.2.2 Các cách thức học	47
3.2.3 Quá trình học nhiều lần của một số BAM.....	47
3.3 Lý do đề xuất thuật toán học mới.....	49
3.4 Thuật toán học mới cho BAM.....	50
3.4.1 Ý tưởng.....	50
3.4.2 Phân tích mối quan hệ giữa MNTP và hàm năng lượng	51
3.4.3 Nội dung thuật toán học mới	52
3.5 Kết quả thực nghiệm	55

3.5.1 Thử nghiệm với nhận dạng vân tay	55
3.5.2 Thử nghiệm với nhận dạng chữ viết tay	56
3.5.3 Thử nghiệm với các ứng dụng nhận dạng khác.....	57
2.6 Kết luận chương	58
CHƯƠNG 4. HAI LUẬT HỌC CẢI TIẾN CHO LÝ THUYẾT CỘNG HƯỞNG THÍCH	
NGHI MỒ 60	
4.1 Giới thiệu chung.....	60
4.2 Các nghiên cứu liên quan	60
4.2.1 Mô hình ART	60
4.2.2 Mô hình Fuzzy ART.....	61
4.2.3 Các luật học điển hình của ART và Fuzzy ART.....	64
4.3 Lý do đề xuất hai luật học	65
4.4 Hai luật học đề xuất cho Fuzzy ART	65
4.4.1 Ý tưởng.....	65
4.4.2 Nội dung của hai luật học	65
4.4.3 Ưu điểm của hai luật học.....	67
4.5 Kết quả thực nghiệm	68
4.5.1 Thử nghiệm 1: Dùng luật học thứ nhất.....	69
4.5.2 Thử nghiệm 2: Dùng luật học thứ hai.....	75
4.6 Kết luận chương	81
CHƯƠNG 5. LUẬT HỌC CẢI TIẾN CHO BỘ NHỚ LIÊN KẾT MỒ.....	82
5.1 Giới thiệu chung.....	82
5.2 Các nghiên cứu liên quan	82
5.2.1 Các mô hình lý thuyết.....	82
5.2.2 Các biến thể của FAM.....	83
5.2.3 Một số mô hình FAM.....	84
5.3 Lý do đề xuất luật học cải tiến cho FAM.....	88
5.4 Luật học cải tiến	88
5.4.1 Ý tưởng.....	88
5.4.2 Mô hình FAM với luật học cải tiến	88
5.4.3 Định lý và hệ quả về khả năng nhớ lại hoàn hảo của FAM cải tiến.....	90
3.5 Kết quả thực nghiệm	91
3.5.1 Thử nghiệm với tập dữ liệu về các số.....	92

5.5.2 Thử nghiệm với tập dữ liệu của Corel.....	93
3.6 Kết luận chương.....	95
KẾT LUẬN.....	97
DANH MỤC CÁC CÔNG TRÌNH KHOA HỌC CỦA TÁC GIẢ LIÊN QUAN ĐẾN LUẬN ÁN.....	99
TÀI LIỆU THAM KHẢO	100

Danh mục các từ viết tắt

Từ viết tắt	Nội dung tiếng Anh	Nội dung tiếng Việt
<i>ACAM</i>	Association-Content Associative Memory	Bộ nhớ liên kết nội dung-liên kết
<i>AM</i>	Associative Memory	Bộ nhớ liên kết
<i>ANN</i>	Artificial Neural Network	Mạng nơ-ron nhân tạo
<i>ART</i>	Adaptive Resonance Theory	Lý thuyết cộng hưởng thích nghi
<i>ARTMAP</i>	Adaptive Resonance Theory Map	Ánh xạ lý thuyết cộng hưởng thích nghi
<i>BAM</i>	Bidirectional Associative Memory	Bộ nhớ liên kết hai chiều
<i>FAM</i>	Fuzzy Associative Memory	Bộ nhớ liên kết mờ
<i>FFBAM</i>	Fast Flexible Bidirectional Associative Memory	Bộ nhớ liên kết hai chiều nhanh-linh động
<i>Fuzzy ART</i>	Fuzzy Adaptive Resonance Theory	Lý thuyết cộng hưởng thích nghi mờ
<i>Fuzzy ARTMAP</i>	Fuzzy Adaptive Resonance Theory Map	Ánh xạ lý thuyết cộng hưởng thích nghi mờ
<i>IFAM</i>	Implicative Fuzzy Associative Memory	Bộ nhớ liên kết mờ gợi ý
<i>MAM</i>	Morphological Associative Memories	Các bộ nhớ liên kết hình thái
<i>MNTP</i>	The Minimum Number of Times for training pairs of Patterns	Số lần tối thiểu cần huấn luyện một cặp mẫu trong Bộ nhớ liên kết hai chiều

Danh mục các bảng

Bảng 3.1: Thời gian học và kết quả nhớ lại các vân tay	56
Bảng 3.2: Thời gian học và kết quả nhớ lại các chữ viết tay.....	56
Bảng 3.3: Thời gian học và kết quả nhớ lại các biển hiệu giao thông.....	57
Bảng 3.4: Thời gian học và kết quả nhớ lại các tiền xu của Mỹ	57
Bảng 3.5: Thời gian học và kết quả nhớ lại các phương tiện giao thông	58
Bảng 4.1: Đặc trưng của các tập dữ liệu trong thử nghiệm 1	69
Bảng 4.2: Kết quả phân lớp đúng của tập Iris.....	70
Bảng 4.3: Kết quả phân lớp đúng của tập Spiral	70
Bảng 4.4: Kết quả phân lớp đúng của tập Flame	71
Bảng 4.5: Kết quả phân lớp đúng của tập Blance-Scale.....	72
Bảng 4.6: Kết quả phân lớp đúng của tập R15	72
Bảng 4.7: Kết quả phân lớp đúng của tập Glass	73
Bảng 4.8: Kết quả phân lớp đúng của tập Wine	73
Bảng 4.9: Kết quả phân lớp đúng của tập Jain	74
Bảng 4.10: Kết quả phân lớp đúng của tập Aggregation.....	74
Bảng 4.11: Sự cải thiện khả năng phân lớp của EFART với luật học thứ nhất so với mô hình tốt nhất thứ hai	75
Bảng 4.12: Đặc trưng của các tập dữ liệu trong thử nghiệm 2	76
Bảng 4.13: Kết quả phân lớp đúng của tập WDBC	77
Bảng 4.14: Kết quả phân lớp đúng của tập D31	77
Bảng 4.15: Kết quả phân lớp đúng của tập WINE-WHITE	77
Bảng 4.16: Kết quả phân lớp đúng của tập BALANCE-SCALE.....	79
Bảng 4.17: Kết quả phân lớp đúng của tập R15	79

Bảng 4.18: Kết quả phân lớp đúng của tập MONK.....	79
Bảng 4.19: Kết quả phân lớp đúng của tập WINE-RED	80
Bảng 4.20: Sự cải thiện khả năng phân lớp của EFART so với mô hình tốt nhất thứ hai trong thử nghiệm 2.....	80
Bảng 5.1: Kết quả của thử nghiệm bộ nhớ tự liên kết với tập dữ liệu về con số.....	92
Bảng 5.2: Kết quả thử nghiệm của bộ nhớ liên kết khác loại với tập dữ liệu về con số	93
Bảng 5.3: Kết quả của thử nghiệm bộ nhớ tự liên kết với tập dữ liệu của Corel	94
Bảng 5.4: Kết quả của thử nghiệm bộ nhớ liên kết khác loại với tập dữ liệu của Corel	94

Danh mục các hình vẽ, đồ thị

Hình 1.1: Cấu tạo nơ-ron sinh học.....	17
Hình 1.2: Mô hình một nơ-ron nhân tạo	18
Hình 1.3: Một mạng truyền thẳng một lớp	20
Hình 1.4: Một mạng truyền thẳng nhiều lớp.....	20
Hình 1.5: Các mạng lan truyền ngược	21
Hình 2.1: Một bộ nhớ nội dung-địa chỉ	29
Hình 2.2: Hai dạng liên kết của bộ nhớ liên kết. Hình 2.2(a) Bộ nhớ dạng tự liên kết. Hình 2.2(b) Bộ nhớ dạng liên kết khác loại.....	30
Hình 2.3: Mô hình mạng Hopfield.....	31
Hình 2.4: Cấu trúc tổng quát của mô hình BAM.....	34
Hình 2.5: Cấu trúc của một ART đơn giản.....	39
Hình 3.1: Các ảnh huấn luyện trong nhận dạng vân tay.	55
Hình 3.2: Các ảnh huấn luyện trong nhận dạng chữ viết tay.....	56
Hình 3.3: Các ảnh huấn luyện trong các ứng dụng nhận dạng khác. Hình 3.3(a) – Các ảnh mẫu về biển hiệu giao thông, Hình 3.3(b) – Các ảnh mẫu về tiền xu của Mỹ, và Hình 3.3(c) – Các ảnh mẫu về các phương tiện giao thông	57
Hình 5.1: Thử nghiệm bộ nhớ tự liên kết với tập dữ liệu về số. Hàng đầu tiên chứa các ảnh huấn luyện gốc; Hàng thứ hai chứa các mẫu vào nhiễu bị biến đổi; Hàng thứ 3,4,5,6 chứa mẫu ra từ mô hình của Junbo, Xiao, Sussner và Valle và ACAM.	92
Hình 5.2: Một số ảnh từ tập dữ liệu của Corel dùng cho thử nghiệm	94
Hình 5.3: Các mẫu thử nghiệm được sinh ra từ các mẫu vào bằng nhiễu muối tiêu	94
Hình 5.4: Các mẫu từ tập dữ liệu của Corel được mô hình đưa ra phục hồi mẫu từ nhiễu muối tiêu tốt hơn các mô hình khác trong chế độ tự liên kết. Từ trái sang phải là các mẫu được phục hồi bởi mô hình của Junbo, Kosko, Xiao, Sussner và Valle, ACAM, và kết quả mong đợi.	95

Hình 5.5: Từ trái sang phải là các mẫu từ tập dữ liệu của Corel được phục hồi từ nhiễu muối tiêu trong chế độ liên kết khác loại bởi các mô hình của Junbo, Kosko, Xiao, Sussner và Valle, ACAM, và kết quả mong đợi.95

Danh mục các thuật toán

Thuật toán 3.1: Thuật toán học nhanh và linh động cho BAM	53
Thuật toán 4.1: Tìm giá trị thích hợp cho tham số tốc độ học của Fuzzy ART.....	66

Danh mục các định lý

Định lý 5.1: (Khả năng nhớ lại các cặp mẫu trong chế độ liên kết khác loại).....	90
Hệ quả 5.1: (Khả năng nhớ lại các cặp mẫu trong chế độ tự liên kết).....	91

MỞ ĐẦU

Mạng nơ-ron nhân tạo (ANN - *Artificial Neural Network*) là tập hợp các đơn vị xử lý thông tin mô phỏng dựa trên hoạt động của hệ thống nơ-ron sinh học. Mỗi ANN có hai quá trình chính gồm quá trình học và quá trình xử lý. Trong quá trình học, ANN thực hiện học và lưu trữ thông tin các dữ liệu mẫu. Trong quá trình xử lý, ANN dùng thông tin học được từ quá trình học để đưa ra tín hiệu ra từ các tín hiệu vào mới. Do đó, chất lượng của quá trình học ảnh hưởng lớn đến chất lượng của quá trình xử lý. Nói cách khác, kết quả ra của quá trình học ảnh hưởng đến hiệu quả xử lý thông tin của mạng. Vì vậy, việc nâng cao chất lượng của quá trình học là hết sức cần thiết để đáp ứng được các yêu cầu xử lý ngày càng phức tạp của các ứng dụng thực.

ANN thường lưu trữ các thông tin học được trong các trọng số kết nối giữa các nơ-ron. Do đó, quá trình học thực hiện cập nhật trọng số kết nối theo một quy tắc được gọi là luật học. Một số luật học điển hình gồm luật lỗi-sửa lỗi, luật Boltzmann, luật Hebb, và luật cạnh tranh. Do kết quả của quá trình học đóng vai trò quyết định đến chất lượng xử lý của ANN nên việc nâng cao chất lượng của quá trình học sẽ làm tăng khả năng xử lý của ANN.

Các ANN có thể chia thành hai nhóm dựa vào mục đích: các ANN phục hồi các mẫu hoàn chỉnh từ các mẫu bị hỏng hóc, và các ANN thực hiện phân loại. Lớp con các ANN thực hiện ghi nhớ các mẫu hoàn chỉnh để có thể nhớ lại các mẫu này từ các mẫu vào hỏng hóc được gọi là ANN ghi nhớ. Với mong muốn nâng cao hiệu quả xử lý của các ANN ghi nhớ, tác giả đã chọn đề tài luận án với tiêu đề: “**Cải tiến quá trình học của một số mạng nơ-ron ghi nhớ**”. Tác giả đã nghiên cứu ba mô hình lý thuyết điển hình và mới nhất trong số các ANN ghi nhớ gồm (i) Bộ nhớ liên kết hai chiều (BAM – *Bidirectional Associative Memory*), (ii) Bộ nhớ liên kết mờ (FAM – *Fuzzy Associative Memory*), và (iii) Lý thuyết cộng hưởng thích nghi mờ (Fuzzy ART – *Fuzzy Adaptive Resonance Theory*). Mục tiêu nghiên cứu của luận án là thực hiện cải tiến quá trình học để nâng cao chất lượng xử lý của ba mô hình trên.

Các nghiên cứu về BAM, FAM, Fuzzy ART đã được tác giả nghiên cứu sâu đặc biệt là quá trình học của mỗi mô hình. Tác giả nhận thấy ba mô hình này đã được phát triển với một số ưu điểm và được áp dụng cho hàng nghìn ứng dụng thực. Tuy nhiên, nếu thực hiện cải tiến thêm quá trình học thì sẽ nâng cao chất lượng xử lý của mỗi mô hình. Lý do đề xuất các cải tiến cho các mô hình gồm:

- Việc gán trọng số cho các mẫu trong tập huấn luyện còn chưa hiệu quả. Với BAM học nhiều lần, có hai cách thể hiện gồm gán trọng số cho các cặp mẫu hoặc học lặp đi lặp lại các cặp mẫu. Tuy nhiên, việc xác định trọng số cho mỗi cặp mẫu còn chưa cụ thể hoặc phải tốn nhiều thời gian điều chỉnh trọng số kết nối [54, 66, 67, 68, 69, 70, 76]. Đối với Fuzzy ART, véc tơ trọng số của các cụm được học để gần giống với các mẫu vào. Tuy nhiên, một số mẫu huấn luyện có thể không được lưu lại trong khi các mẫu này có thể chứa các thông tin quan trọng [8, 9, 38, 41, 61, 73].
- Luật học chưa học hiệu quả các mẫu huấn luyện. Do một số FAM dùng tỷ lệ giá trị tương ứng của mẫu vào và mẫu ra để lưu trữ nên nội dung của các mẫu chưa được quan tâm [58, 65, 71]. Ngược lại, một số FAM chỉ lưu mẫu vào hay mẫu ra nên việc lưu sự liên kết của cặp mẫu lại giảm [14, 40, 58]. Kết quả là mỗi mô hình chỉ thích hợp với một số tập mẫu nhất định nên khó thích hợp với một ứng dụng phải xử lý với các mẫu phức tạp.

Với mong muốn đóng góp một phần vào nâng cao chất lượng xử lý của ba mô hình trên, luận án của tác giả tập trung vào ba hướng sau:

1. Đề xuất cách xác định giá trị thích hợp cho trọng số của các cặp mẫu huấn luyện của BAM
2. Cải tiến luật học và tìm giá trị thích hợp cho tham số học của Fuzzy ART.
3. Cải tiến luật học để FAM học và lưu trữ hiệu quả đồng thời cả nội dung và sự liên kết giữa các cặp mẫu.

Các kết quả của luận án gồm: 2 bài báo công bố ở Tạp chí quốc tế có chỉ số ISI [18, 27], 7 báo cáo được công bố trong kỳ yếu của các hội nghị quốc tế có phần

biện (trong đó, 6 báo cáo được xuất bản bởi IEEE và Springer) [4, 5, 28, 29, 30, 31, 33], và 2 bài báo ở tạp chí trong nước [32, 34] .

Nội dung luận án được tổ chức thành 5 chương. Chương đầu giới thiệu các khái niệm cơ bản về ANN. Tiếp theo, cơ sở lý thuyết được trình bày trong Chương 2. Các đóng góp của luận án được trình bày trong ba chương tiếp theo. Ngoài ra, các phần mở đầu, mục lục, kết luận, tài liệu tham khảo cũng được viết ở phần đầu và cuối của luận án. Nội dung của từng chương được trình bày như sau:

Chương 1 trình bày các kiến thức quan trọng về ANN gồm nơ-ron sinh học, mô hình nơ-ron nhân tạo, mạng nơ-ron, các luật học, ưu-nhược điểm, và ứng dụng. Các kiến thức này giúp hiểu chủ đề nghiên cứu của luận án.

Chương 2 cung cấp các kiến thức cơ bản về các khái niệm, thao tác của logic mờ và toán học hình thái. Tiếp theo, mô hình và hoạt động của BAM, FAM, và Fuzzy ART được trình bày chi tiết hơn giúp phân biệt điểm mới trong các đề xuất cải tiến đối với từng mô hình.

Chương 3 đề xuất một thuật toán xác định trọng số thích hợp cho mỗi cặp mẫu huấn luyện của BAM học nhiều lần. BAM cải tiến được thử nghiệm với ứng dụng nhận dạng mẫu. Kết quả thực nghiệm cho thấy có khả năng nhớ lại được cải thiện hơn các BAM khác. Thuật toán học cải tiến giúp BAM học nhanh và linh động hơn nhưng vẫn đảm bảo về khả năng phục hồi mẫu.

Chương 4 thể hiện hai luật học và một thủ tục tìm giá trị thích hợp cho tham số tốc độ học của Fuzzy ART. Các thực nghiệm phân cụm trên 14 bộ dữ liệu chuẩn cho thấy Fuzzy ART với các luật học đề xuất nhớ lại các cụm tốt hơn. Trong chương, cả luật học và tham số mô hình đều được cải tiến để nâng cao chất lượng phân cụm Fuzzy ART.

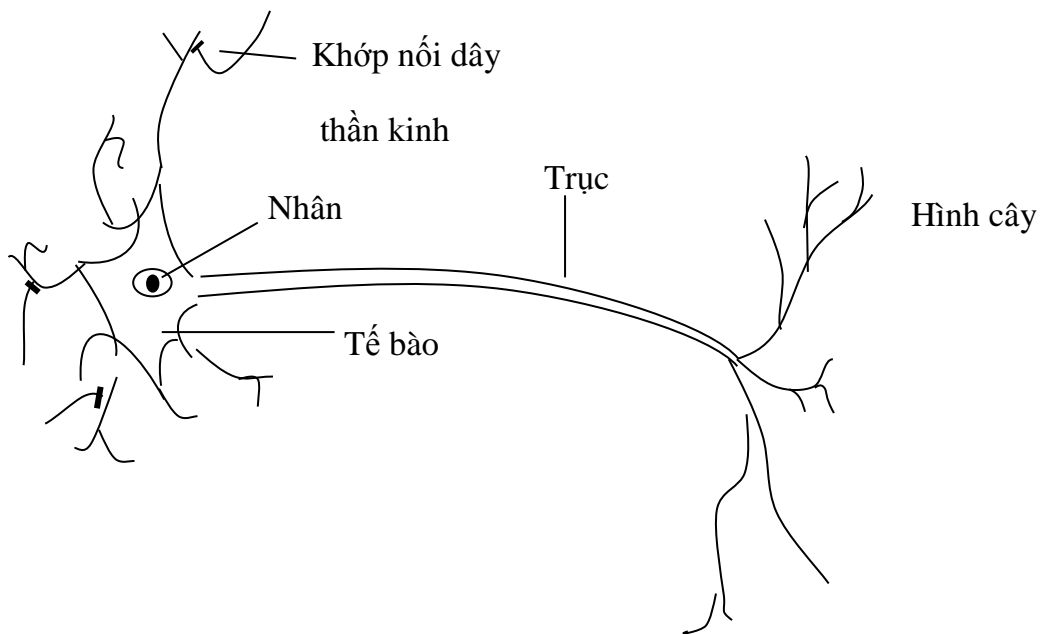
Chương 5 trình bày luật học cho FAM. Các thử nghiệm với nhiệm vụ nhận dạng mẫu cho thấy FAM với luật học cải tiến nhớ lại tốt hơn các FAM khác. Luật học đề xuất đã giúp nâng cao khả năng phục hồi mẫu từ các mẫu vào có dạng nhiễu phức tạp.

CHƯƠNG 1. MẠNG NƠ-RON NHÂN TẠO

Trong chương này, tác giả trình bày các kiến thức cơ bản về mạng nơ-ron nhân tạo.

1.1 Nơ-ron sinh học

Một nơ-ron sinh học [26] là một tế bào xử lý và truyền thông tin bằng các tín hiệu hóa học qua một khớp thần kinh tới các tế bào khác. Mỗi nơ-ron kết nối với nơ-ron khác hình thành các mạng nơ-ron.



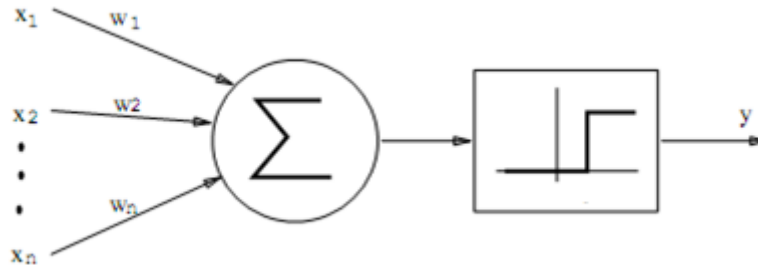
Hình 1.1

Hình 1.1: Cấu tạo nơ-ron sinh học.

Một nơ-ron có một thân tế bào, các hình cây và một sợi trục như trong Hình 1.1. Các hình mọc ra từ thân tế bào và chia thành nhiều nhánh. Một sợi trục được sinh ra từ thân tế bào. Các tín hiệu được truyền đi từ một sợi trục của một nơ-ron tới một hình cây của nơ-ron khác.

1.2 Nơ-ron nhân tạo

McCulloch và Pitts [16] đã đề xuất một mô hình tính toán cho một nơ-ron sinh học như trong Hình 1.2.



Hình 1.2: Mô hình một nơ-ron nhân tạo

Trong đó:

- y là tín hiệu ra
- x_1, x_2, \dots, x_n là các tín hiệu vào
- w_1, w_2, \dots, w_n là các trọng số tương ứng với các tín hiệu vào

Hoạt động của nơ-ron gồm hai bước: tổng hợp các tín hiệu vào và tạo ra một tín hiệu ra dựa vào các tín hiệu vào.

Tổng hợp các thông tin vào

Cho các thông tin vào x_1, x_2, \dots, x_n với các trọng số tương ứng w_1, w_2, \dots, w_n được thực hiện theo một trong các công thức sau:

Dạng tuyến tính

$$input = \sum_{j=1}^n w_j x_j \quad (1.1)$$

Dạng toàn phương

$$input = \sum_{j=1}^n w_j x_j^2 \quad (1.2)$$

Dạng mặt cầu

$$input = \frac{1}{\rho^2} \sum_{j=1}^n (w_j - x_j)^2 \quad (1.3)$$

với ρ là bán kính của mặt cầu

Tạo ra tín hiệu ra

ANN dùng một hàm kích hoạt để xác định tín hiệu ra. Dưới đây là một số dạng hàm thường được dùng.

Hàm McCulloch-Pitts:

$$out = f(input) = \begin{cases} 1 & \text{nếu } input \geq \theta \\ 0 & \text{nếu } input < \theta \end{cases} \quad (1.4)$$

với θ là ngưỡng.

Hàm McCulloch-Pitts trễ:

$$out = f(input) = \begin{cases} 1 & \text{if } input \geq UTP \\ 0 & \text{if } input < LTP \\ f(input) & \text{otherwise} \end{cases} \quad (1.5)$$

với $UTP > LTP$ và UTP là ngưỡng trên, còn LTP là ngưỡng dưới

Hàm sigmoid

$$out = f(input) = \frac{1}{1 + e^{-\lambda \cdot input}} \quad (1.6)$$

với λ là độ nghiêng của hàm

1.3 Mạng nơ-ron nhân tạo

Mạng nơ-ron nhân tạo [60] là một cấu trúc được hình thành do các nơ-ron nhân tạo liên kết với nhau. Mỗi nơ-ron có các tín hiệu vào, tín hiệu ra và thực hiện một chức năng tính toán cục bộ.

Các đặc điểm nổi bật của ANN gồm:

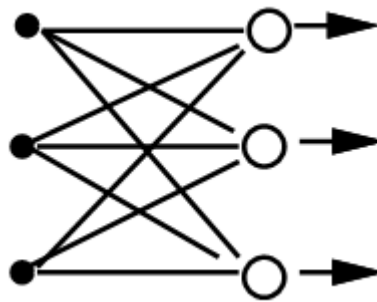
- Là mô hình toán học dựa trên bản chất hoạt động của nơ-ron sinh học

- Cấu tạo từ một số các nơ-ron có liên kết với nhau
- Có khả năng học và tổng quát hóa tập dữ liệu thông qua việc gán và hiệu chỉnh trọng số liên kết giữa các nơ-ron
- Xử lý song song các thông tin phân tán nên có khả năng tính toán lớn

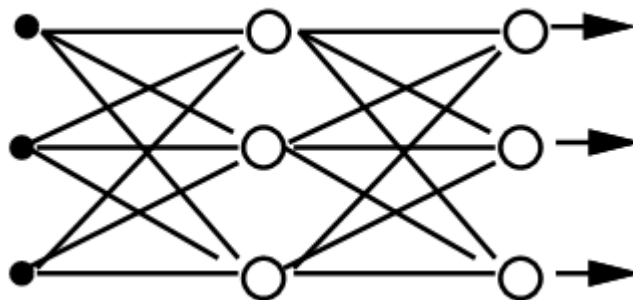
Các ANN có thể chia theo nhiều cách dựa vào cấu trúc (một tầng và nhiều tầng), cách truyền tín hiệu (truyền thẳng và lan truyền ngược), và bản chất của việc học (học giám sát, học không giám sát, học lai giữa 2 cách)

Phân loại theo cách truyền tín hiệu

Mạng truyền thẳng: gồm các mạng perceptron một lớp, mạng perceptron nhiều tầng và mạng RBF. Hình 1.3 và Hình 1.4 thể hiện một mạng truyền thẳng một lớp và nhiều lớp.

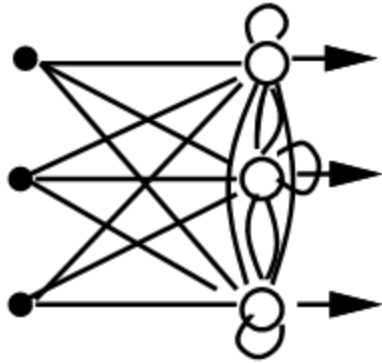


Hình 1.3: Một mạng truyền thẳng một lớp



Hình 1.4: Một mạng truyền thẳng nhiều lớp

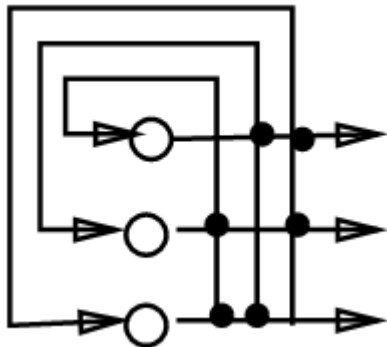
Mạng lan truyền ngược: gồm các mạng cạnh tranh, mạng SOM của Kohonen, mạng Hopfield và mô hình ART. Hình 1.5 mô tả các loại mạng này.



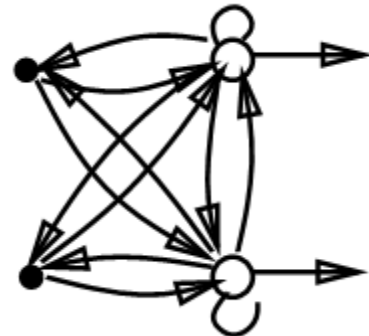
(a) Mạng cạnh tranh



(b) Mạng SOM của Kohonen



(c) Mạng Hopfield



(d) Mạng ART

Hình 1.5: Các mạng lan truyền ngược

Phân loại theo cách học dữ liệu

Mạng học giám sát: mạng được cung cấp tín hiệu ra đúng của mỗi tín hiệu vào. Các trọng số liên kết được xác định để tạo ra tín hiệu ra giống nhất với tín hiệu

ra đúng. Học tăng cường là trường hợp đặc biệt của học có giám sát, do mạng chỉ được cung cấp điều kiện về tín hiệu ra đúng.

Mạng học không giám sát: Tìm ra các cấu trúc ẩn của dữ liệu, sự tương quan giữa các mẫu, và tổ chức các mẫu thành các nhóm dựa vào sự tương quan.

Học lai là một phần trọng số thu được nhờ học giám sát và phần còn lại thu được nhờ học không giám sát.

1.4 Các luật học của ANN

Các luật học đóng vai trò quan trọng trong việc xác định một ANN. Quá trình học của ANN [1] là cập nhật ma trận trọng số, các tham số mô hình dựa vào các mẫu huấn luyện. Theo nghĩa rộng thì học có thể chia làm hai loại: Học tham số và học cấu trúc.

Học tham số: Các thủ tục học này tìm kiếm ma trận trọng số sao cho mạng có khả năng đưa ra các dự báo sát với thực tế. Dạng chung của luật học tham số cho nơ-ron i được mô tả như sau:

$$\Delta W_{ij} = \eta r x_j, \text{ với } j = 1..n \quad (1.7)$$

Trong đó:

- ΔW_{ij} là sự thay trọng số liên kết của nơ-ron thứ i do nơ-ron j tạo ra
- x_j là tín hiệu vào nơ-ron j .
- η là tốc độ học, nằm trong khoảng $(0,1)$.
- r là hằng số học.

Có thể chia thủ tục học tham số ra thành hai lớp nhỏ hơn gồm học có chỉ đạo và học không chỉ đạo. Việc xác định r phụ thuộc vào từng kiểu học.

+ *Học có tín hiệu chỉ đạo:* dựa vào sai số của tín hiệu ra thực và tín hiệu ra mong muốn để hiệu chỉnh trọng số. Sai số này chính là hằng số học r . Luật điển hình của nhóm là luật học Delta dựa trên nguyên tắc giảm gradient. Tiếp đến là luật học perceptron, luật học OJA, và luật lan truyền ngược cho mạng nhiều lớp.

+ *Học không có tín hiệu chỉ đạo*: sử dụng tín hiệu ra của mạng làm cơ sở để hiệu chỉnh các trọng số liên kết. Điển hình là luật Hebb thường dùng cho các mạng tự liên kết, luật LVQ (Learning Vector Quantization) dùng cho mạng tự tổ chức một lớp thuộc lớp mạng ánh xạ đặc trưng của Kohonen.

Luật học Hebb dựa trên hiện tượng sinh học sau: Giữa hai nơ-ron có quan hệ và có thay đổi thế năng màng thì giữa chúng có sự thay đổi trọng số liên kết. Nói cách khác, trọng số được điều chỉnh theo mối tương quan giữa tín hiệu vào và tín hiệu ra.

$$\Delta W_{ij} = \eta y_i x_j, \text{ với } j = 1..n \quad (1.8)$$

Trong đó:

- ΔW_{ij} là sự thay đổi trọng số liên kết của nơ-ron i do các nơ-ron j tạo ra
- x_j là tín hiệu vào nơ-ron j.
- y_i là tín hiệu ra của nơ-ron i.
- η là tốc độ học, nằm trong khoảng (0,1).

Luật Hebb giải thích việc điều chỉnh trọng số trong phạm vi cục bộ của mạng khi không cần tín hiệu chỉ đạo từ bên ngoài. Hopfield cũng cải tiến luật Hebb cho các mạng tự liên kết thành 16 dạng khác nhau của luật Hebb, luật đối Hebb, luật Hopfield.

Đối với mạng lan truyền ngược thường sử dụng luật Hebb và các luật cải tiến của luật Hebb để điều chỉnh trọng số.

+ *Học tăng cường*: Trong một số trường hợp, thông tin phản hồi chỉ là tín hiệu bao gồm hai trạng thái cho biết tín hiệu đầu ra của mạng là đúng hay sai. Quá trình học dựa trên các thông tin hướng dẫn như vậy được gọi là học có củng cố (học tăng cường) và tín hiệu mang thông tin phản hồi được gọi là tín hiệu củng cố cho quá trình học. Đây là một dạng của học có tín hiệu chỉ đạo.

Học cấu trúc: Tìm kiếm các tham số về cấu trúc để có được một mạng hoạt động tốt nhất. Trong thực tế, việc học cấu trúc là tìm ra số lớp ẩn và tìm ra số nơ-ron trên mỗi lớp đó. Giải thuật di truyền thường được sử dụng trong các cấu trúc

nhưng thường chạy rất lâu. Ngoài ra, kỹ thuật gọt tĩa mạng hay mạng tăng dần cũng được áp dụng trong việc học cấu trúc của mạng có kích thước nhỏ.

1.5 Ưu và nhược điểm của ANN

Ưu điểm

- Xử lý song song.
- Thiết kế hệ thống thích nghi.
- Không đòi hỏi các đặc trưng mở rộng của bài toán (chủ yếu dựa trên tập mẫu học).

Nhược điểm

- Không có các quy tắc và các hướng dẫn thiết kế một cách rõ ràng đối với một ứng dụng nhất định.
- Không có cách tổng quát để đánh giá hoạt động bên trong mạng.
- Việc học đối với mạng có thể khó (hoặc không thể) thực hiện.
- Khó có thể dự đoán trước được hiệu quả của mạng trong tương lai (khả năng tổng quát hoá).

1.6 Ứng dụng của ANN

Mạng nơ-ron được coi như là hộp đen biến đổi véc-tơ đầu vào m biến thành véc-tơ đầu ra n biến. Tín hiệu ra có thể là các số thực, (tốt nhất nằm trong khoảng $[0,1]$, hoặc $[-1,1]$), số nhị phân $0,1$, hay số lưỡng cực $-1,+1$. Số biến của véc-tơ vào/ véc-tơ ra không bị hạn chế xong sẽ ảnh hưởng tới thời gian tính toán và tải dữ liệu của máy tính. Nói chung, các lớp bài toán áp dụng cho nơ-ron có thể được phân chia thành bốn loại gồm phân lớp, mô hình hoá, biến đổi, (thực hiện ánh xạ từ một không gian đa biến vào không gian đa biến khác tương ứng), và liên kết và kỹ thuật dịch chuyển cửa sổ.

Phân lớp

Để quản lý các đối tượng thường phân lớp đối tượng vào các nhóm, nhóm con, hay chủng loại. Ví dụ: Bài toán phân lớp ảnh, nhận dạng mẫu, . . .

Khi việc phân lớp phức tạp, cần nghiên cứu thống kê các mối liên quan giữa nhiều đối tượng và thuộc tính của lớp các đối tượng. Việc xây dựng một cây phân lớp được thực hiện trong quá trình học. Nếu kết quả phân loại không tốt, cần xét lại cách biểu diễn các đối tượng hoặc cây phân lớp hoặc thay đổi cả hai.

Mô hình hoá

Mục đích của mô hình hóa là có thể đưa ra các dự báo cho tất cả các đối tượng đầu vào. Việc tìm ra đường cong phù hợp với các số liệu thực nghiệm là một trong những ứng dụng thuộc dạng này. Mô hình cũng phải tuân theo một giả định: các thay đổi nhỏ của tín hiệu vào chỉ gây ra những biến đổi nhỏ của tín hiệu ra.

Trong nhiều ứng dụng với sai số thực hiện khá lớn, có thể mô hình hoá bằng cách cân đối giữa tín hiệu vào và tín hiệu ra. Khi đó, mạng được sử dụng như một bảng tra mặc dù các bảng này sẽ cho lời giải giống nhau trong một khoảng giá trị của tín hiệu vào.

Biến đổi

Việc biến đổi nhằm mục đích nén các đối tượng từ không gian nhiều chiều vào không gian có số chiều nhỏ hơn rất nhiều. Qua việc nén, các đối tượng này sẽ bộc lộ các đặc điểm không thể nhận thấy khi chúng thuộc không gian nhiều chiều. Biến đổi tương tự như việc nhóm hay phân lớp các đối tượng thể hiện ở chỗ biểu diễn các kết quả ra. Điểm quan trọng trong biến đổi là các đối tượng được biểu diễn bởi tọa độ của nơ-ron trung tâm chứ không phải là giá trị của tín hiệu ra.

Một trong những ứng dụng của việc biến đổi là tiền xử lý để chọn ra các đối tượng điển hình từ tập vô số các đối tượng ngẫu nhiên, loại trừ các đối tượng dư thừa hay trùng lặp. Điều này là cực kỳ quan trọng khi lựa chọn các đối tượng làm mẫu học cho mạng lan truyền ngược có dùng sai số.

Liên kết

Liên kết là tìm ra đối tượng đích có mối quan hệ với một đối tượng vào ngay cả khi đối tượng vào bị hỏng hoặc hoàn toàn không biết.

Kỹ thuật dịch chuyển cửa sổ

Nghiên cứu quá trình phụ thuộc thời gian là lĩnh vực chính trong nghiên cứu quá trình điều khiển do có thể dự báo được các hành vi của hệ thống dựa trên một chuỗi số liệu được ghi nhận theo thời gian.

Việc học dịch chuyển tới bước tiếp theo tạo ra các cửa sổ bao gồm số bước thời gian của véc-tơ ra. Kỹ thuật dịch chuyển cửa sổ có thể được sử dụng để giải quyết các vấn đề chuỗi các sự kiện và đối tượng như trong các lĩnh vực về môi trường theo thời gian, kiểm soát hỏng hóc.

1.7 Kết luận chương

Trong chương này, các hiểu biết cơ bản về ANN được trình bày bao gồm mô hình nơ-ron nhân tạo, mạng nơ-ron nhân tạo, và các luật học. Các kiến thức này cung cấp cách nhìn tổng quát về chủ đề nghiên cứu của luận án.

CHƯƠNG 2. MỘT SỐ MẠNG NƠ-RON GHI NHỚ

Trong chương này, tác giả trình bày các kiến thức toán học làm cơ sở mô tả thiết kế các mô hình ANN gồm Logic mờ, Toán học hình thái. Sau đó, tác giả trình bày các ANN được phát triển trong luận án gồm Bộ nhớ liên kết, Bộ nhớ liên kết hai chiều, Bộ nhớ liên kết mờ, Lý thuyết cộng hưởng thích nghi, và Lý thuyết cộng hưởng thích nghi mờ.

2.1 Logic mờ

Logic mờ [74] được phát triển để thể hiện các khái niệm mờ, mơ hồ trong ngôn ngữ tự nhiên khi biểu diễn các thông tin không chính xác.

2.1.1 Định nghĩa

Cho X là một miền xác định. Một tập mờ A trên miền X được đặc trưng bằng một hàm thuộc

$$\mu_A: X \rightarrow [0,1]$$

với $\mu_A(x)$ thể hiện mức độ thuộc của phần tử x vào tập mờ A đối với mọi $x \in X$.

2.1.2 Các phép toán với tập mờ

Phép nối mờ của A, B là một ánh xạ:

$$C: [0,1] \times [0,1] \rightarrow [0,1] \quad (2.1)$$

thỏa mãn $C(0, 0) = C(0, 1) = C(1, 0) = 0$ và $C(1, 1) = 1$. Ví dụ: phép lấy min và tích.

Một kết nối mờ

$$T: [0,1] \times [0,1] \rightarrow [0,1] \quad (2.2)$$

thỏa mãn $T(x, 1) = x$ với $x \in [0,1]$ được gọi là *triangular norm* hay *t-norm*.

Các kết nối mờ C_M, C_P và C_L là các *t-norm*.

$$C_M(x, y) = x \wedge y \quad (2.3)$$

$$C_P(x, y) = x \cdot y \quad (2.4)$$

$$C_L(x, y) = 0 \vee (x + y - 1) \quad (2.5)$$

với \vee là phép lấy giá trị lớn nhất và \wedge là phép lấy giá trị nhỏ nhất

Một phân tách mờ là một ánh xạ

$$D: [0,1] \times [0,1] \rightarrow [0,1] \quad (2.6)$$

thỏa mãn $D(0, 0) = 0$ và $D(0, 1) = D(1, 0) = D(1, 1) = 1$.

Một phân tách mờ

$$S: [0,1] \times [0,1] \rightarrow [0,1] \quad (2.7)$$

thỏa mãn $S(1, x) = x$ với $x \in [0, 1]$ được gọi là *triangular co-norm* hay *short s-norm*.

Các thao tác sau biểu diễn các *s-norm*:

$$D_M(x, y) = x \vee y \quad (2.8)$$

$$D_P(x, y) = x + y - x \cdot y \quad (2.9)$$

$$D_L(x, y) = 1 \wedge (x + y) \quad (2.10)$$

với \wedge là phép lấy giá trị nhỏ nhất

Một thao tác

$$I: [0, 1] \times [0, 1] \rightarrow [0, 1] \quad (2.11)$$

được gọi là gọi ý mờ nếu I mở rộng suy diễn thông thường $[0, 1] \times [0, 1]$ với $I(0, 0) = I(0, 1) = I(1, 1) = 1$ và $I(1, 0) = 0$.

Một vài gọi ý mờ cụ thể:

$$I_M(x, y) = \begin{cases} 1, & x \leq y \\ y, & x > y \end{cases} \quad (2.12)$$

$$I_P(x, y) = \begin{cases} 1, & x \leq y \\ y/x, & x > y \end{cases} \quad (2.13)$$

$$I_L(x, y) = 1 \wedge (y - x + 1) \quad (2.14)$$

với \wedge là phép lấy giá trị nhỏ nhất

2.2 Toán học hình thái

Toán học hình thái [56] là một lý thuyết tập trung vào xử lý và phân tích các đối tượng bằng việc sử dụng các thao tác và chức năng dựa trên hình dạng và các khái niệm hình học. Hầu hết các kết quả toán học của toán học hình thái được thể hiện trên các lưới đầy đủ

2.2.1 Lưới đầy đủ

Lưới đầy đủ là một tập có thứ tự, từng phần \mathbb{L} trong đó mỗi tập con có một chặn trên và chặn dưới trong \mathbb{L} . Với mỗi $Y \subseteq \mathbb{L}$, chặn dưới của Y được ký hiệu $\wedge Y$ và chặn trên được ký hiệu $\vee Y$. Lớp các tập mờ kế thừa cấu trúc lưới đầy đủ với tập giá trị thuộc miền $[0, 1]$.

2.2.2 Các thao tác cơ bản với lưới đầy đủ

Phép co rút là một ánh xạ ε từ một lưới đầy đủ \mathbb{L} đến một lưới đầy đủ \mathbb{M} thỏa mãn công thức sau:

$$\varepsilon\left(\bigwedge Y\right) = \bigwedge_{y \in Y} \varepsilon(y) \quad (2.15)$$

Phép giãn nở $\delta: \mathbb{L} \rightarrow \mathbb{M}$ thỏa mãn công thức sau:

$$\delta\left(\bigvee Y\right) = \bigvee_{y \in Y} \delta(y) \quad (2.16)$$

2.3 Mô hình AM

2.3.1 Khái niệm về AM

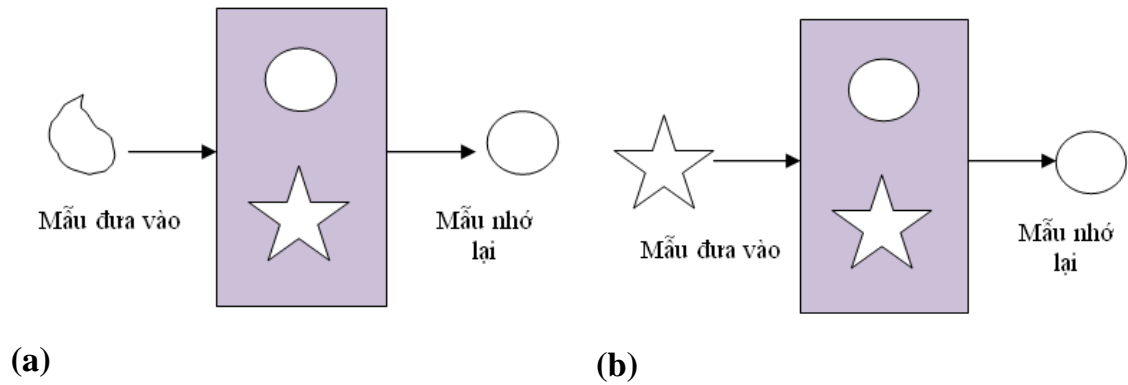
AM [36] là một cấu trúc nội dung-địa chỉ thực hiện ánh xạ các mẫu vào sang các mẫu ra. AM là một dạng bộ nhớ cho phép nhớ lại mẫu đã lưu dựa vào mức độ tương tự giữa mẫu vào và các mẫu đã lưu. Hình 2.1 mô tả một ví dụ về bản chất của AM. Khi đưa vào một mẫu vào nhiều hay không chính xác thì AM dựa vào các mẫu đã lưu để tìm ra mẫu giống với mẫu vào nhất để làm mẫu ra. Đây là một dạng làm đúng các lỗi.



Hình 2.1: Một bộ nhớ nội dung-địa chỉ

2.3.2 Hoạt động của AM

AM có hai dạng liên kết gồm tự liên kết và liên kết khác loại. Bộ nhớ ở dạng tự liên kết đưa ra một mẫu đã lưu giống nhất với mẫu vào hiện tại. Ở dạng liên kết khác loại, mẫu ra khác hoàn toàn mẫu vào về nội dung, kiểu và định dạng nhưng có liên quan với mẫu vào. Hình 2.2 mô tả hai dạng AM.



Hình 2.2: Hai dạng liên kết của bộ nhớ liên kết. Hình 2.2(a) Bộ nhớ dạng tự liên kết. Hình 2.2(b) Bộ nhớ dạng liên kết khác loại

Cặp mẫu trực giao là hai mẫu được biểu diễn dưới dạng hai véc tơ một chiều có tích vô hướng bằng 0.

AM có khả năng đưa ra mẫu ra đúng từ một tập các mẫu vào nhiều hoặc không đầy đủ trong cả dạng tự liên kết và liên kết khác loại. Do đó, khi bộ nhớ được kích hoạt với một mẫu vào thì mẫu lưu trữ trong bộ nhớ được nhớ lại (xuất hiện ở đầu ra). Mẫu vào có thể chính xác, nhiều hay là biểu diễn từng phần của mẫu được lưu trong bộ nhớ.

AM có hai quá trình gồm quá trình học và quá trình nhớ lại. Với quá trình học, các cặp mẫu được lưu trong ma trận trọng số kết nối. Quá trình nhớ lại thực hiện phục hồi một mẫu đã lưu từ các mẫu vào hồng học thông qua sự nhớ lại các mẫu đã lưu trong ma trận trọng số. Do đó, quá trình học và nhớ lại liên quan mật thiết với nhau.

2.3.3 Một số đặc điểm của AM

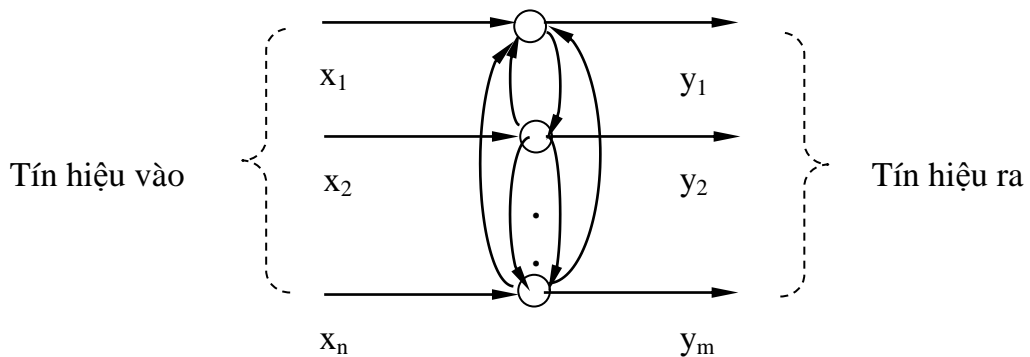
- Các lỗi và nhiễu chỉ gây ra giảm độ chính xác của mẫu ra và không ảnh hưởng đến sự thực hiện của mạng
- Nếu các cặp mẫu được mã hóa thành cặp véc tơ trực giao thì AM nhớ lại đúng cặp mẫu đó. Ngược lại, AM không thể nhớ lại do sự đan chéo giữa các mẫu trong bộ nhớ.

- Các mẫu thường mã hóa thành véc tơ với các giá trị ở dạng 2 cực để kiểm tra được tính trực giao của mỗi cặp mẫu.

2.4 Mô hình BAM

2.4.1 Mạng Hopfield

Mạng Hopfield [35] là mô hình tiêu biểu của lớp mạng lan truyền ngược. Mạng Hopfield là mạng một lớp có rất nhiều ứng dụng, đặc biệt trong bộ nhớ liên kết và trong các bài toán tối ưu. Hình 2.3 mô tả mô hình mạng Hopfield.



Hình 2.3: Mô hình mạng Hopfield

Tín hiệu ra của nơ-ron thứ j nào đó được truyền ngược lại làm tín hiệu vào cho các nơ-ron thông qua các trọng số tương ứng.

Ký hiệu w_{ij} là trọng số liên kết giữa hai nơ-ron i và j ($w_{ij} = w_{ji}$), y_i là đầu ra của nơ-ron i . Khi đó, véc tơ (y_1, y_2, \dots, y_n) là trạng thái của mạng. Tại mỗi thời điểm t mỗi nơ-ron i tổng hợp các tín hiệu x_j từ các nơ-ron khác và tín hiệu từ bên ngoài I_i

$$input_i = \sum_{j=1}^n w_{ij}x_j + I_i \quad (2.17)$$

Tuỳ theo hàm kích hoạt f_i , nơ-ron i cho tín hiệu ra.

$$y_i(t+1) = f_i(y_i(t)) \quad (2.18)$$

Mạng đạt trạng thái cân bằng nếu $y_i(t+1) = y_i(t), \forall i$

Hàm năng lượng của mạng được tính bằng:

$$E = E(y_1, \dots, y_n) = -\frac{1}{2} \sum_{i=1}^n \sum_j^n w_{ij} y_i y_j - \sum_{i=1}^n I_i y_i \quad (2.19)$$

Tùy theo phương thức hoạt động, có thể chia thành mạng Hopfield rời rạc và mạng Hopfield liên tục.

Mạng Hopfield rời rạc

Mạng Hopfield rời rạc [60] là mạng có tín hiệu ra là rời rạc và làm việc ở chế độ không đồng bộ. Tín hiệu ra nhận các giá trị nhị phân $\{0, 1\}$:

Hàm kích hoạt được xác định như sau:

$$y_i = f(\text{input}_i) = \begin{cases} 1 & \text{if } \text{input}_i \geq 0 \\ 0 & \text{if } \text{input}_i < 0 \end{cases} \quad (2.20)$$

Việc cho hàm kích hoạt trên tương đương với quy tắc chuyển trạng thái $y_i(t+1) = y_i(t) + \Delta y_i$ trong đó Δy_i được cho bởi công thức:

$$\Delta y_i = \begin{cases} 1 & \text{if } \sum_{j=1}^n w_j x_j + I_i > 0 \text{ and } y_i(t) = 0 \\ -1 & \text{if } \sum_{j=1}^n w_j x_j + I_i \leq 0 \text{ and } y_i(t) = 1 \\ 0, & \text{otherwise} \end{cases} \quad (2.21)$$

Định lý: Giả sử $W_{ii}=0, i=1, \dots, n$. Khi đó, với quy tắc chuyển trạng thái trên và cập nhật không đồng bộ thì năng lượng của mạng không tăng (tức là giảm hoặc giữ nguyên)

Chứng minh: Giả sử nơ-ron k thay đổi trạng thái từ thời điểm t đến $t+1$. Khi đó mạng sẽ thay đổi năng lượng và

$$\Delta E = E(t+1) - E(t) = \left(- \sum_j w_{kj} y_j(t) \right) \Delta y_k$$

vì thế theo công thức tính Δy_i luôn có $\Delta E \leq 0$, tức là năng lượng của mạng không tăng. Vì thế, hàm năng lượng sẽ đạt tới giá trị cực tiểu do hàm giới nội.

Do tính chất hội tụ và giá trị nhị phân của các nơ-ron nên mạng Hopfield rời rạc được sử dụng cho các bài toán tối ưu $\{0, 1\}$

Mạng Hopfield liên tục

Mạng Hopfield liên tục [60] là mạng có trạng thái được mô tả bởi phương trình động học sau:

$$\frac{d(input_i)}{dt} = \sum_{j=1}^n w_j x_j + I_i \quad (2.22)$$

và $y_i = f(input_i)$ với f_i là hàm kích hoạt

Giả sử $W_{ij}=W_{ji}$ và $W_{ii}=0$. Nếu hàm năng lượng được cho bởi công thức (2.22) thì :

$$\frac{d(input_i)}{dt} = \frac{\partial E}{\partial y_i} \quad (2.23)$$

Sự hội tụ của mạng Hopfield liên tục cho bởi định lý sau:

Định lý: Nếu $f_i(input_i)$ ($i=1, \dots, n$) là các hàm khả vi và không giảm thì $\frac{dE}{dt} \leq 0$

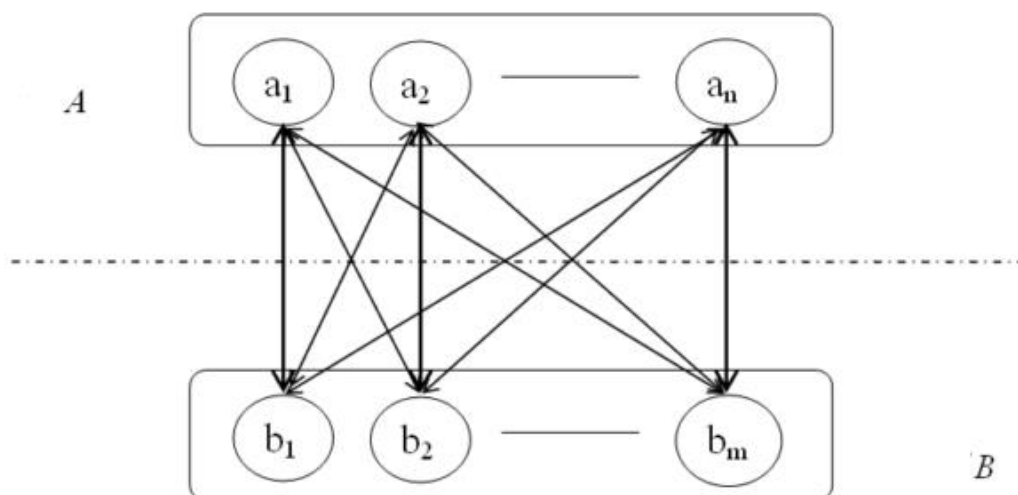
Chứng minh: Ta có

$$\frac{dE}{dt} = \sum_j \frac{\partial E}{\partial y_j} \frac{dy_j}{dinput_j} \frac{dinput_j}{dt} = - \sum_j \left(\frac{\partial E}{\partial y_j} \right)^2 \cdot \frac{dy_j}{dinput_j}$$

vì theo giả thiết các hàm $f_i(input_i)$ là không giảm nếu $\frac{dy_j}{dinput_j} \geq 0$ do đó $\frac{dE}{dt} \leq 0$

2.4 2 Khái niệm về BAM

BAM [45] là một AM thể hiện cấu trúc bộ nhớ liên kết với khả năng nhớ lại theo cả hai hướng. BAM được cấu tạo từ hai mạng Hopfield để thực hiện liên kết giữa hai mẫu. BAM cũng có hai dạng gồm tự liên kết (khi mẫu vào và mẫu ra trong một cặp là giống nhau) và liên kết khác loại (khi mẫu vào và mẫu ra trong một cặp là khác nhau). Hình 2.4 mô tả cấu trúc tổng quát của mô hình BAM.



Hình 2.4: Cấu trúc tổng quát của mô hình BAM

Trong mô hình này, BAM lưu p liên kết khác loại giữa hai trường A và B. Các cặp mẫu được kí hiệu là: $(\mathbf{A}^1, \mathbf{B}^1)$, ..., $(\mathbf{A}^p, \mathbf{B}^p)$. Khi cung cấp mẫu vào từ trường A thì BAM sẽ nhớ lại mẫu đã lưu ở trường B. Ngược lại, cung cấp mẫu vào từ trường B thì thu được mẫu ra ở trường A. Các mẫu ở hai trường A, B phải được biểu diễn thành véc tơ chứa các giá trị ở dạng hai cực gồm hai giá trị 0 và 1 hay -1 và 1 để tiện kiểm tra tính trực giao của cặp mẫu.

2.4.3 Quá trình học của BAM

Quá trình học thực hiện học sự liên kết giữa các cặp mẫu. Sau đó, tổng quát hóa các liên kết và lưu trữ trong một ma trận trọng số chung.

Quá trình học được thực hiện như sau:

Đầu tiên, ma trận trọng số \mathbf{W}^k lưu liên kết của cặp mẫu $(\mathbf{A}^k, \mathbf{B}^k)$ được tính theo công thức sau:

$$\mathbf{W}^k = \mathbf{A}^k \mathbf{B}^{kT} \quad (2.24)$$

với \mathbf{A}^k là ma trận cấp $1 \times n$, \mathbf{B}^k là ma trận cấp $1 \times m$, và \mathbf{W}^k là ma trận cấp $n \times m$.

Sau đó, tổng quát hóa sự liên kết của p cặp mẫu và lưu trong ma trận trọng số chung, \mathbf{W} – Ma trận trọng số gốc.

$$\mathbf{W} = \sum_{k=1}^p \mathbf{W}^k \quad (2.25)$$

2.4.4 Quá trình nhớ lại của BAM

Quá trình nhớ lại thực hiện đưa ra một mẫu đã lưu có liên quan đến mẫu vào. Cho một mẫu vào \mathbf{X} , quá trình nhớ lại diễn ra như sau:

Đầu tiên, tổng hợp tín hiệu vào của mỗi nơ-ron theo công thức sau:

$$Input_j = \sum_{i=1}^n X_i W_{ij} \quad (2.26)$$

với

n là số chiều của mẫu vào \mathbf{X}

$Input_j$ là tổng các tín hiệu vào của nơ-ron j

X_i là thành phần thứ i của \mathbf{X}

Sau đó, xác định tín hiệu ra cho nơ-ron bằng cách dùng hàm đầu ra:

$$Y_j = \begin{cases} 1, & \text{if } Input_j \geq 0 \\ -1, & \text{if } Input_j < 0 \end{cases} \quad (2.27)$$

Tiếp tục, \mathbf{Y} là mẫu vào của BAM từ phía B, lặp lại quy trình tính toán trên với hai công thức (2.29) và (2.30) nhưng ma trận trọng số chung \mathbf{W} từ hướng B sang A bằng chuyển vị của ma trận trọng số chung từ hướng A sang B. Kết quả ra ký hiệu là \mathbf{X}^1 . Sau đó, \mathbf{X}^1 lại được xem là mẫu vào của BAM và thu được \mathbf{Y}^1 .

Lặp lại quá trình trên cho đến khi thu được cặp $(\mathbf{X}^f, \mathbf{Y}^f)$ không thay đổi. Đây là trạng thái BAM hội tụ và \mathbf{Y}^f chính là mẫu ra của BAM ứng với mẫu vào \mathbf{X} .

2.4.5 Hàm năng lượng của BAM

Hàm năng lượng (hàm Lyapunov) là một hàm gắn với mỗi trạng thái của BAM. Mỗi trạng thái được biểu diễn bằng một cặp mẫu. Hàm có tính chất là giảm dần theo thời gian.

Để lưu và nhớ lại được một cặp mẫu thì hàm năng lượng phải đạt đến một cực tiểu cục bộ và không được phá hủy các cặp mẫu đã lưu.

Hàm năng lượng E^k với cặp mẫu $(\mathbf{A}^k, \mathbf{B}^k)$.

$$E^k(\mathbf{A}^k, \mathbf{B}^k) = -\mathbf{A}^k \mathbf{W} \mathbf{B}^{kT} \quad (2.28)$$

Đưa vào cặp $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ để thu được cặp gần nhất với $(\mathbf{A}^i, \mathbf{B}^i)$, các nơ-ron phải thay đổi cho đến khi mạng ổn định với cặp mẫu $(\mathbf{A}^f, \mathbf{B}^f)$.

Kosko đã chứng minh BAM chỉ hội tụ khi hàm năng lượng đạt cực tiểu cục bộ. Do đó, nếu năng lượng ứng với cặp mẫu $(\mathbf{A}^i, \mathbf{B}^i)$ không đạt cực tiểu cục bộ thì không thể nhớ lại ngay cả khi $\boldsymbol{\alpha} = \mathbf{A}^i$.

2.4.6 Chiến lược học nhiều lần dùng số lần lặp tối thiểu để học một cặp mẫu

Y.F. Wang và đồng nghiệp [69] đưa ra mô hình BAM thực hiện học nhiều lần để đảm bảo nhớ lại đúng các cặp mẫu đã lưu. Khi đó ma trận trọng số \mathbf{W}^k lưu cặp mẫu $(\mathbf{A}^k, \mathbf{B}^k)$ được tính theo công thức:

$$\mathbf{W}^k = q^k \mathbf{A}^k \mathbf{B}^{kT} \quad (2.29)$$

với q^k là số dương thể hiện số lần ít nhất dùng $(\mathbf{A}^k, \mathbf{B}^k)$ cho việc học để đảm bảo nhớ lại được $(\mathbf{A}^k, \mathbf{B}^k)$. q^k được viết tắt là MNTTP.

2.5 Mô hình FAM

2.5.1 Khái niệm FAM

AM lưu sự liên kết của các cặp mẫu có liên quan và có khả năng nhớ lại các mẫu đã lưu. AM được mô tả như sau:

Cho một tập các liên kết $(\mathbf{A}^k, \mathbf{B}^k)$, $k=1, \dots, p$ xác định một ánh xạ \mathbf{G} sao cho $\mathbf{G}(\mathbf{A}^k) = \mathbf{B}^k$ với mọi $k=1, \dots, p$. Hơn nữa, ánh xạ \mathbf{G} cần có khả năng chịu nhiễu. Nghĩa là, $\mathbf{G}(\mathbf{A}'^k)$ nên bằng \mathbf{B}^k đối với các bản nhiễu hay không đầy đủ \mathbf{A}'^k của \mathbf{A}^k .

Tập các liên kết $(\mathbf{A}^k, \mathbf{B}^k)$, $k=1, \dots, p$ được gọi là tập bộ nhớ cơ bản và mỗi liên kết $(\mathbf{A}^k, \mathbf{B}^k)$ trong tập này được gọi là bộ nhớ cơ bản [36]. Một bộ nhớ tự liên kết là tập bộ nhớ cơ bản với dạng $(\mathbf{A}^k, \mathbf{A}^k)$, $k=1, \dots, p$. Bộ nhớ được gọi là liên kết khác loại nếu mẫu ra \mathbf{B}^k là khác với mẫu vào \mathbf{A}^k .

Quá trình xác định \mathbf{G} được gọi là quá trình học và ánh xạ \mathbf{G} thực hiện nhớ lại các liên kết.

Bộ nhớ liên kết mờ là bộ nhớ liên kết với các mẫu \mathbf{A}^k và \mathbf{B}^k là các tập mờ với mọi $k=1, \dots, p$.

2.5.2 Các kiểu nơ-ron trong FAM

Pedrycz [50] đưa ra lớp các nơ-ron mờ tổng quát nhất do các nơ-ron này tổng quát hóa một nhóm các mẫu vào và các trọng số liên kết.

Giả sử, \mathbf{W} là ma trận lưu các trọng số liên kết, n là số phần tử của véc tơ biểu diễn mẫu vào và θ là sai số.

Nơ-ron Max-C

Đây là mô hình nơ-ron được dùng phổ biến nhất. Với \mathbf{x} là mẫu vào, mẫu ra \mathbf{y} được nhớ lại theo cách sau:

$$\mathbf{y} = \left[\bigvee_{j=1}^n C(\mathbf{W}_j, \mathbf{x}_j) \right] \vee \theta \quad (2.30)$$

với $C()$ là phép nối mờ của logic mờ ở dạng t-norm.

Nơ-ron Min-I

Mẫu ra \mathbf{y} được nhớ lại từ mẫu vào \mathbf{x} được tính như sau:

$$\mathbf{y} = \left[\bigwedge_{j=1}^n I(\mathbf{W}_j, \mathbf{x}_j) \right] \wedge \theta \quad (2.31)$$

với $I()$ là phép gợi ý mờ của logic mờ.

Nơ-ron Min-D

Cho \mathbf{x} là mẫu vào, mẫu ra \mathbf{y} được nhớ lại theo cách sau:

$$\mathbf{y} = \left[\bigwedge_{j=1}^n D(\mathbf{W}_j, \mathbf{x}_j) \right] \wedge \theta \quad (2.32)$$

với D() là phép phân tách mờ của logic mờ ở dạng s-norm.

2.5.3 Các FAM của Kosko và sự tổng quát hóa

Kosko [43, 44] đưa ra hai mô hình FAM đầu tiên gồm max-min FAM và max-product FAM. Sau đó, Chung và Lee [12] tổng quát hóa FAM thành FAM tổng quát.

Giả sử, FAM lưu p cặp mẫu. Cho $\mathbf{X} = [\mathbf{X}^1, \dots, \mathbf{X}^p] \in [0,1]^{n \times p}$ và $\mathbf{Y} = [\mathbf{Y}^1, \dots, \mathbf{Y}^p] \in [0,1]^{m \times p}$

Max-min FAM

Mô hình này dùng nơ-ron max- C_M . Quá trình học thực hiện theo công thức sau:

$$W_{ij} = \bigvee_{k=1}^p C_M(W_{ik}, x_{kj}), i = 1..m, j = 1..n \quad (2.33)$$

Với x là mẫu vào, mẫu ra y được nhớ lại theo cách sau:

$$\mathbf{y} = \left[\bigvee_{j=1}^n C_M(\mathbf{W}_j, \mathbf{x}_j) \right] \quad (2.34)$$

với $C_M(x, y) = x \wedge y$

Max-Product FAM

Mô hình này dùng nơ-ron max- C_P . Quá trình học thực hiện theo công thức sau:

$$W_{ij} = \bigvee_{k=1}^p C_P(W_{ik}, x_{kj}), i = 1..m, j = 1..n \quad (2.35)$$

Với x là mẫu vào, mẫu ra y được nhớ lại theo cách sau:

$$\mathbf{y} = \left[\bigvee_{j=1}^n C_P(\mathbf{W}_j, \mathbf{x}_j) \right] \quad (2.36)$$

với $C_p(x, y) = x \cdot y$

FAM tổng quát

Mô hình tổng dùng nơ-ron max-C nên có thể dùng một phép nối mờ như C_M , C_P , C_L .

Quá trình học thực hiện theo công thức sau:

$$W_{ij} = \bigvee_{k=1}^p C(W_{ik}, x_{kj}), i = 1..m, j = 1..n \quad (2.37)$$

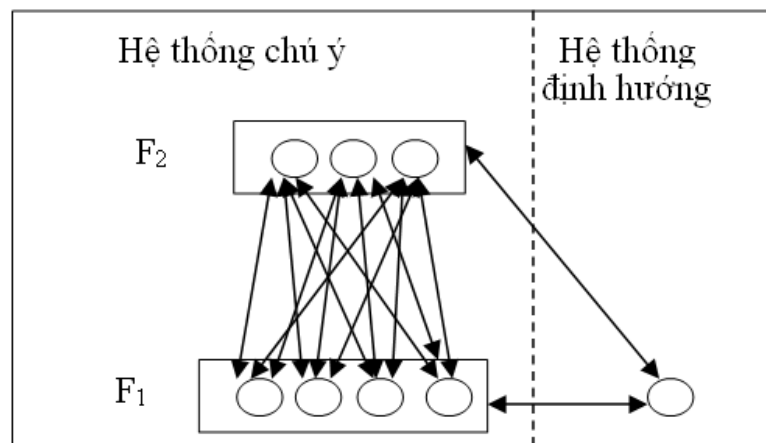
Với x là mẫu vào, mẫu ra y được tính như sau:

$$y = \left[\bigvee_{j=1}^n C(W_j, x_j) \right] \quad (2.38)$$

2.6 Mô hình ART

2.6.1 Cấu trúc của ART

Các ART [24,25] được phát triển bởi Grossberg để giải quyết vấn đề về hiện tượng ổn định-thay đổi. Cấu trúc chung của mạng ART được thể hiện trong Hình 2.5.



Hình 2.5: Cấu trúc của một ART đơn giản

Một mạng ART điển hình có hai tầng: tầng dữ liệu vào (F1) và tầng dữ liệu ra (F2). Tầng dữ liệu vào chứa N nút với N là số lượng các mẫu vào. Số lượng nút của tầng dữ liệu ra là động. Mỗi nút của tầng dữ liệu ra có một véc tơ kiểu tương ứng với mỗi cụm.

Tính động của mạng được điều khiển bởi hai hệ thống con: hệ thống chú ý và hệ thống định hướng. Hệ thống chú ý đưa ra một nơ-ron chiến thắng (cụm) và hệ thống định hướng quyết định cụm nào chấp nhận hay không chấp nhận mẫu vào đó.

Mạng ART ở trạng thái cộng hưởng khi hệ thống định hướng chấp nhận một cụm chiến thắng khi véc tơ kiểu của cụm chiến thắng khớp đủ gần với mẫu vào hiện tại.

2.6.2 Các bước hoạt động chính của ART

Hoạt động của ART gồm 3 bước chính: chọn một cụm chiến thắng, kiểm tra điều kiện về trạng thái cộng hưởng, và học mẫu huấn luyện.

Các mẫu vào và véc tơ trọng số của các cụm được biểu diễn thành các véc tơ có giá trị được thể hiện ở dạng nhị phân.

ART sử dụng hai tham số gồm tham số chọn α và tham số ngưỡng $\rho \in [0,1]$ (điều kiện để một cụm chấp nhận mẫu huấn luyện hiện tại).

Mỗi cụm j có một véc tơ trọng số của cụm, $\mathbf{W}_j = (W_{j1}, \dots, W_{jM})$.

Ký hiệu \cap là thao tác logic AND. Nghĩa là, $x_i \cap y_i = 1$ nếu $x_i = 1$ và $y_i = 1$, còn các trường hợp còn lại $x_i \cap y_i = 0$.

Chọn một cụm chiến thắng

Với mỗi mẫu vào \mathbf{I} và cụm j , hàm chọn T_j được định nghĩa bởi

$$T_j(\mathbf{I}) = \frac{\|\mathbf{I} \cap \mathbf{W}_j\|}{\alpha + \|\mathbf{W}_j\|} \quad (2.39)$$

và với chuẩn $\|\cdot\|$ được định nghĩa:

$$\|\mathbf{x}\| = \sum_{k=1}^n x_k \quad (2.40)$$

Chọn cụm J với $T_j = \max\{T_j, j = 1..N\}$

Kiểm tra trạng thái cộng hưởng

Cộng hưởng xuất hiện nếu hàm đối chiếu của cụm được chọn đạt điều kiện về ngưỡng:

$$\frac{\|I \cap W_j\|}{\|I\|} \geq \rho \quad (2.41)$$

Học nhanh mẫu huấn luyện

Véc tơ trọng số của cụm thứ J , W_j , được cập nhật theo công thức sau:

$$W_j^{\text{new}} = I \cap W_j^{\text{old}} \quad (2.42)$$

2.6.3 Họ các mô hình của ART

Họ các mô hình được phát triển từ ART để giải quyết tốt hơn các bài toán thực. Đầu tiên, ART 1 có thể học ổn định để phân loại các mẫu vào nhị phân. Thứ hai là ART 2 với khả năng phân loại mẫu vào tương tự hoặc nhị phân. Tiếp theo là ART 3 với khả năng tìm kiếm song song trong hệ thống phân cấp mạng nhiều mức.

Sau này, ART 1 được dùng để thiết kế các kiến trúc mạng phân cấp được gọi là ARTMAP. Các mạng này có thể tự tổ chức nhanh các ánh xạ phân lớp giữa các véc tơ thể hiện mẫu vào và các véc tơ thể hiện mẫu ra có n chiều. Thông qua học có giám sát, ARTMAP tạo ra các lớp nhận dạng có kích thước tối ưu bằng cách cực đại sự tổng quát hóa dự đoán trong khi cực tiểu lỗi dự đoán. ARTMAP được dùng để học các ánh xạ giữa véc tơ thể hiện mẫu ra và véc tơ thể hiện mẫu vào nhị phân.

2.7 Mô hình Fuzzy ART

2.7.1 So sánh với ART

Fuzzy ART được phát triển từ ART 1 để có khả năng học các lớp nhận dạng với cả mẫu vào tương tự và mẫu vào nhị phân. Do đó, Fuzzy ART thay phép AND trong logic bằng phép giao của logic mờ trong các bước hoạt động.

Mọi bước trong thuật toán xử lý của Fuzzy ART giống với ART. Do đó, Fuzzy ART vẫn có được các đặc trưng và ưu điểm của ART. Điểm khác là thay các

thao tác logic AND bằng phép giao (\wedge) của logic mờ và mẫu được thể hiện bằng tập mờ. Hơn nữa, Fuzzy ART có thêm tham số tốc độ học β trong việc cập nhật trọng số.

2.7.2 Thuật toán Fuzzy ART

Thuật toán này được Carpenter trình bày ngắn gọn trong [9]. Ba tham số sau thể hiện tính động của mô hình Fuzzy ART:

- Tham số chọn $\alpha > 0$;
- Tham số tốc độ học $\beta \in [0, 1]$
- Tham số ngưỡng $\rho \in [0, 1]$;

Nội dung của thuật toán được trình bày như sau:

Bước 1: Khởi tạo véc tơ trọng số của các cụm tiềm năng.

Mỗi cụm j có một véc tơ trọng số $\mathbf{W}_j = (W_{j1}, \dots, W_{jM})$. Số các cụm tiềm năng N là bất kỳ. Khởi tạo

$$W_{j1} = \dots = W_{jM} = 1 \quad (2.43)$$

và mỗi cụm được coi là chưa hình thành. Sau khi một cụm được chọn để mã hóa, cụm được hình thành. W_{ji} là không tăng dần theo thời gian nên các W_{ji} hội tụ tới một giới hạn.

Bước 2: Lựa chọn một cụm chiến thắng.

Với mỗi mẫu vào \mathbf{I} và cụm j , hàm chọn T_j được định nghĩa bởi

$$T_j(\mathbf{I}) = \frac{\|\mathbf{I} \wedge \mathbf{W}_j\|}{\alpha + \|\mathbf{W}_j\|} \quad (2.44)$$

với phép toán giao, \wedge , trong logic mờ được định nghĩa:

$$(x \wedge y)_i = \min\{x_i, y_i\} \quad (2.45)$$

và với chuẩn $\|\cdot\|$ được định nghĩa như trong Công thức 2.40.

Để đơn giản, $T_j(\mathbf{I})$ được viết là T_j khi mẫu vào \mathbf{I} cố định. Sự chọn cụm được gán chỉ số bằng J , với

$$T_j = \max\{T_j, j = 1..N\} \quad (2.46)$$

Nếu có nhiều hơn một T_j là cực đại thì cụm j với chỉ số nhỏ nhất được chọn.

Bước 3: Kiểm tra trạng thái của mạng là cộng hưởng hay thiết lập lại.

Cộng hưởng xuất hiện nếu hàm đối chiếu của cụm được chọn đạt điều kiện về ngưỡng:

$$\frac{\|\mathbf{I} \wedge \mathbf{W}_j\|}{\|\mathbf{I}\|} \geq \rho \quad (2.47)$$

Sau đó, việc học sẽ diễn ra.

Thiết lập lại xuất hiện nếu

$$\frac{\|\mathbf{I} \wedge \mathbf{W}_j\|}{\|\mathbf{I}\|} < \rho \quad (2.48)$$

Sau đó, giá trị của hàm chọn T_j được thiết lập -1 để ngăn sự chọn lại cụm J trong quá trình tìm kiếm. Một chỉ số mới J được chọn bởi Công thức (2.46). Quá trình tìm kiếm tiếp tục cho đến khi J được chọn thỏa mãn Công thức (2.47). Nếu không có cụm đang tồn tại nào thỏa mãn điều kiện trong Công thức (2.47) thì một cụm mới J được sinh ra và đặt $\mathbf{W}_j^{\text{new}} = \mathbf{I}$.

Bước 4: Học dữ liệu huấn luyện.

Véc tơ trọng số của cụm thứ J , \mathbf{W}_j được cập nhật theo công thức sau:

$$\mathbf{W}_j^{\text{new}} = \beta (\mathbf{I} \wedge \mathbf{W}_j^{\text{old}}) + (1 - \beta) \mathbf{W}_j^{\text{old}} \quad (2.49)$$

2.7.3 Fuzzy ART với mã hóa đầy đủ

Moore [49] mô tả vấn đề sinh ra cụm mới trong các ART tương tự khi một số lớn các mẫu vào khác biệt lớn so với véc tơ trọng số của các cụm. Việc sinh các cụm được tránh nếu mẫu vào được chuẩn hóa bằng cách chọn $\gamma > 0$ sao cho

$$\|\mathbf{I}\| = \gamma \quad (2.50)$$

với mọi mẫu vào \mathbf{I} . Chuẩn hóa có thể làm được bằng cách tiền xử lý mỗi véc tơ thể hiện mẫu vào \mathbf{a} . Một luật chuẩn hóa, được gọi là mã hóa đầy đủ khi bảo toàn thông tin về độ lớn. Mã hóa đầy đủ biểu diễn cả thông tin thực và thông tin bù của \mathbf{a} . Đặt

\mathbf{a} biểu diễn các thông tin thực. Phần bù của \mathbf{a} , được ký hiệu bởi \mathbf{a}^c , biểu diễn phần thông tin bù với

$$\mathbf{a}_i^c = \mathbf{1} - \mathbf{a}_i \quad (2.51)$$

2.7.3 Thước đo chất lượng phân cụm

Hai thước đo cơ bản được dùng phổ biến cho phân cụm gồm

Davies–Bouldin index

Giá trị của chỉ số này được tính như sau [22]:

$$DB = \frac{1}{n} \sum_{i=1}^n \max_{i \neq j} \left(\frac{\sigma_i + \sigma_j}{d(c_i, c_j)} \right) \quad (2.52)$$

với n là số cụm, \mathbf{c}_x là trung tâm của cụm x , σ_x là khoảng cách trung bình từ tất cả các phần tử của cụm x tới trung tâm cụm \mathbf{c}_x , và $d(c_i, c_j)$ là khoảng cách giữa hai trung tâm của cụm i và j .

Thuật toán có chỉ số Davies-Bouldin càng bé càng tốt.

Dunn index

Công thức tính chỉ số này [22] được trình bày như sau:

$$D = \min_{1 \leq i \leq n} \left\{ \min_{1 \leq j \leq n, i \neq j} \left\{ \frac{d(i, j)}{\max_{1 \leq k \leq n} d'(k)} \right\} \right\} \quad (2.53)$$

với $d(i, j)$ là khoảng cách giữa cụm i và cụm j , $d'(k)$ là khoảng cách giữa các phần tử trong cụm k . Có nhiều cách đo $d(i, j)$ khác nhau.

Thuật toán có chỉ số Dunn càng lớn càng tốt.

2.8 Kết luận chương

Trong chương này, tác giả trình bày các kiến thức toán học cơ bản và các mô hình ANN ở dạng bộ nhớ. Các mô hình này gồm mạng Hopfield, Bộ nhớ liên kết, Bộ nhớ liên kết hai chiều, Bộ nhớ liên kết mờ, Lý thuyết cộng hưởng thích nghi, và Lý thuyết cộng hưởng thích nghi mờ.

CHƯƠNG 3. THUẬT TOÁN HỌC CẢI TIẾN CHO BỘ NHỚ LIÊN KẾT HAI CHIỀU

Trong phần 2.3 và 2.4 của chương 2, tác giả đã trình bày các hiểu biết quan trọng về AM và BAM. Trong chương này, tác giả sẽ trình bày các nghiên cứu có liên quan đến BAM để làm cơ sở lý luận đề xuất cải tiến. Tiếp theo, thuật toán học mới cho BAM và các kết quả thực nghiệm sẽ được mô tả và phân tích chi tiết hơn.

3.1 Giới thiệu chung

BAM là một kiểu AM được mở rộng từ mạng Hopfield để thực hiện tìm kiếm sự liên kết theo cả hai chiều. BAM có một ưu điểm là nhớ lại một mẫu đã lưu từ một mẫu vào có chứa nhiều hoặc không đầy đủ. Hơn nữa, BAM hội tụ không điều kiện trong chế độ đồng bộ. Đây là một đặc trưng ưu việt hơn mạng Hopfield và giúp cho BAM có thể áp dụng cho các ứng dụng thực tế.

3.2 Các nghiên cứu liên quan

3.2.1 Các mô hình lý thuyết

Một số mô hình mới được tạo ra để cải tiến khả năng lưu trữ và nhớ lại. Y.F Wang và đồng nghiệp [68,69,70] đưa ra các điều kiện hiệu quả và cần thiết cho trọng số của ma trận tương quan tổng quát. Các điều kiện này đảm bảo BAM nhớ lại mọi cặp mẫu huấn luyện sau khi thực hiện chiến lược huấn luyện nhiều lần. Zhuang và đồng nghiệp [76] phát triển các luật học tốt hơn dựa vào ba điều kiện tối ưu về sự ổn định của các vùng hấp dẫn (các vùng có chứa cực tiểu cục bộ của hàm năng lượng) và có ít nhất các bộ nhớ giả (các vùng của hàm năng lượng gần giống với cực tiểu cục bộ). Nhóm tác giả đã đưa ra khái niệm ổn định Hamming của các vùng hấp dẫn. Do đó, luật học Perceptron của Rosenblatt [53] được dùng để thu được tính ổn định của các vùng hấp dẫn và các điều kiện tối ưu. Xu và He [72] đưa ra mô hình BAM với các kết nối không đối xứng bên trong và khả năng chứa một số lượng lớn các cặp mẫu không trực giao. Hơn nữa, các ưu điểm của BAM vẫn đạt được nhưng không tăng độ phức tạp của mạng. T. Wang và đồng nghiệp [66,67] đưa ra một thuật toán học với sự ổn định tối ưu của các vùng hấp dẫn của BAM.

Luật học đưa ra đảm bảo lưu trữ các mẫu huấn luyện với các vùng hấp dẫn lớn nhất. Hơn nữa, các tác giả còn nghiên cứu khả năng lưu trữ, sự hội tụ của phương pháp học, sự ổn định và vùng hấp dẫn của mỗi mẫu huấn luyện. Leung [46] đưa ra một luật học mới để cải thiện khả năng nhớ lại. BAM của Leung BAM có khả năng lưu trữ tốt hơn và có khả năng làm đúng lỗi tốt hơn BAM của Kosko. Shi và đồng nghiệp [57] đưa ra một mô hình chung không yêu cầu trọng số kết nối trong giữa hai nơ-ron. Nhóm tác giả định nghĩa hàm hỗ trợ để đo mức hỗ trợ của trạng thái này cho các trạng thái khác. Sau đó, hàm hỗ trợ được dùng trong quá trình nhớ và thuật toán học được phát triển dựa vào luật học của Rosenblatt. Eom và đồng nghiệp [19,20] điều chỉnh khoảng cách Hamming trong quá trình nhớ lại của BAM không đối xứng bằng cách tăng khả năng lưu trữ và chịu nhiễu. Shen và Cruz [54] mở rộng BAM bằng cách thực hiện quá trình học dựa vào việc tối ưu hàm năng lượng. Trọng số của ma trận tương quan của các cặp mẫu được xác định để thu được tập chịu nhiễu cực đại. BAM này sẽ nhớ lại đúng nếu mẫu vào nằm trong tập chịu nhiễu cực đại. Các tác giả cũng chứng minh tập chịu nhiễu cực đại là lớn nhất và dùng giải thuật di truyền tính toán các trọng số để làm cực đại hàm mục tiêu. Acevedo-mosqueda và đồng nghiệp [2] đã trình bày một bộ dịch Anh-Tây Ban Nha dựa vào một BAM cho phép nhớ lại các mẫu đã lưu một cách đơn giản. Vázquez và đồng nghiệp [64] cũng đưa ra một BAM mới dựa vào sự mở rộng mô hình liên kết động. Mô hình mới này chỉ lưu liên kết theo chiều xuôi nhưng có thể nhớ lại từ cả hai chiều. Chartier và Boukadoum [10,11] giới thiệu một BAM với một luật học theo thời gian và một hàm đầu ra không tuyến tính. Mô hình này có khả năng học trực tuyến nhưng không bị học quá (*overlearning*) và gây ra ít bộ nhớ hấp dẫn giả hơn.

Các nghiên cứu trên đã đề xuất các luật học nhiều lần theo thời gian để đảm bảo nhớ lại đúng các mẫu đã lưu. Tuy nhiên, độ phức tạp tính toán của quá trình học khá lớn. Ngoài ra, một số ít nghiên cứu đưa ra cách thức học các cặp mẫu chỉ một lần nhưng khả năng nhớ lại từ mẫu vào nhiễu còn hạn chế. Hơn nữa, việc nhớ lại đúng chỉ xảy ra khi cặp mẫu được thể hiện thành cặp véc tơ trực giao.

3.2.2 Các cách thức học

Có hai chiến lược học được phát triển gồm học một lần và học nhiều lần. BAM với học một lần được thực hiện rất nhanh trong một lần lặp duy nhất. Một số mô hình học lần lượt từng cặp mẫu trong một lần lặp với các phép toán cơ bản của ma trận như các mô hình của Zhuang và đồng nghiệp [76], Xu và He [72], và Leung [46]. Eom và đồng nghiệp [20] học lần lượt từng cặp mẫu và ma trận đường chéo của các mẫu trong một lần lặp với các phép toán cơ bản của ma trận. Acevedo-mosqueda và đồng nghiệp [2] đưa ra Alpha-Beta BAM với thao tác nhị phân α, β cùng với hai phép biến đổi véc tơ (mở rộng và rút gọn). Vázquez và đồng nghiệp [64] đưa ra một BAM mới với sự mã hóa các mẫu huấn luyện bằng cách phép toán cơ bản của toán học. Sau đó, các mẫu đã mã hóa được học bằng các phép tính của ma trận và phép lấy phần tử trung gian của dãy số.

Các chiến lược học nhiều lần được đưa ra để cải thiện khả năng nhớ lại (khả năng phục hồi mẫu đã lưu từ các mẫu vào nhiều). Y.F Wang và đồng nghiệp [69] thể hiện chiến lược huấn luyện nhiều lần thông qua MNTP. T. Wang và đồng nghiệp [67] đưa ra thuật toán học có trọng số dựa vào giá trị trung bình của các cực tiểu cục bộ. Tập các mẫu được học lần lượt trong nhiều lần lặp. Shi và đồng nghiệp [57] học nhiều lần các mẫu và các biến thể của mẫu bằng các phép tính với véc tơ. Chartier and Boukadoum [11] đưa ra một thuật toán học ngẫu nhiên từng cặp mẫu cho đến khi ma trận trọng số hội tụ.

Các mô hình thực hiện học một lần có độ phức tạp tính toán nhỏ nhưng khả năng chịu nhiễu lại thấp. Ngược lại, các mô hình thực hiện học nhiều lần có khả năng chịu nhiễu cao nhưng độ phức tạp lại cao.

3.2.3 Quá trình học nhiều lần của một số BAM

Giả sử, BAM lưu liên kết của p cặp mẫu từ vùng A sang vùng B với vùng A có các mẫu $\mathbf{A}^1, \dots, \mathbf{A}^p$ và vùng B có các mẫu $\mathbf{B}^1, \dots, \mathbf{B}^p$. Mỗi mẫu ở vùng A được biểu diễn bằng một ma trận cấp $1 \times n$. Tương tự, mẫu ở vùng B là một ma trận cấp $1 \times m$. Ma trận \mathbf{W} cấp $n \times m$ lưu liên kết của các cặp mẫu từ vùng A sang vùng B.

Mô hình của Y.F Wang, Cruz, và Mulligan [68,69,70]

Mô hình này học lần lượt các mẫu trong một lần lặp nhưng thể hiện chiến lược học nhiều lần do sử dụng MNTP. Luật học được thể hiện bởi công thức sau:

$$W_{ij} = \sum_{k=1}^p q^k A_i^k (B_j^k)^T \quad (3.1)$$

$$q^k \geq \max \left(1, \frac{\epsilon_{0i}^A}{2m} + 1, \frac{\epsilon_{0k}^B}{2n} + 1 \right) \quad (3.2)$$

với ϵ_{0i}^A là sự chênh lệch năng lượng lớn nhất giữa cặp mẫu thứ i và các cặp mẫu khác trong vùng A. Tương tự với ϵ_{0i}^B là sự chênh lệch năng lượng lớn nhất giữa cặp mẫu thứ i và các cặp mẫu khác trong vùng B.

Mô hình của T. Wang và Zhuang [66,67]

Mô hình BAM này học lần lượt các mẫu trong nhiều lần lặp của quá trình học. Luật học của mô hình được thể hiện bởi công thức sau:

$$W_{ij}(t+1) = W_{ij}(t) + \Delta W_{ij} \quad (3.3)$$

với số gia trọng số ΔW_{ij} được tính bởi công thức sau:

$$\Delta W_{ij} = \sum_{k=1}^p A_i^k B_j^k [S(\alpha_i^k) + S(\beta_j^k)] \quad (3.4)$$

với $S(x)=0$ nếu $x>0$ và $S(x)=1$ nếu $x\leq 0$.

Hai công thức sau được dùng để tính α_i^k và β_j^k

$$\alpha_i^k = \left(\sum_{j=1}^n W_{ij} B_j^k \right) A_i^k \quad (3.5)$$

$$\beta_j^k = \left(\sum_{i=1}^m W_{ij} A_i^k \right) B_j^k \quad (3.6)$$

Quá trình này lặp lại cho đến khi các giá trị của ma trận trọng số \mathbf{W} ổn định

Mô hình của Zhuang, Huang, và Chen [76]

Mô hình BAM này học lần lượt các mẫu trong nhiều lần lặp của quá trình học. Ban đầu, $W_{ij}^{(0)}$ là bất kỳ. Khi $t > 0$ thì luật học của nơ-ron i ở vùng A được thể hiện bởi công thức sau:

$$W_{ij}(t+1) = \begin{cases} W_{ij}(t), & \text{if } \mathbf{A}_i \sum_{l=1}^n W_{il}(t) \mathbf{B}_l > 0 \\ W_{ij}(t) + \mathbf{B}_j \mathbf{A}_i, & \text{otherwise} \end{cases} \quad (3.7)$$

với $j = 1, \dots, n$.

Luật học của của nơ-ron j ở vùng B được thể hiện bởi công thức sau:

$$W_{ij}(t+1) = \begin{cases} W_{ij}(t), & \text{if } \mathbf{B}_j \sum_{l=1}^m W_{lj}(t) \mathbf{A}_l > 0 \\ W_{ij}(t) + \mathbf{B}_j \mathbf{A}_i, & \text{otherwise} \end{cases} \quad (3.8)$$

với $i = 1, \dots, m$.

Quá trình này lặp lại cho đến khi các giá trị của ma trận trọng số \mathbf{W} ổn định

Các nghiên cứu trên chủ yếu dựa vào hàm năng lượng của BAM để đề xuất các phương thức học khác nhau. Mỗi mô hình hướng tới việc học các mẫu để đảm bảo khả năng nhớ lại hoàn hảo các mẫu đã lưu. Vì vậy, khả năng nhớ lại của các mô hình học khá tốt nhưng độ phức tạp tính toán lại cao. Ngoài ra, các mô hình này chủ yếu tập trung vào lưu trữ các cặp mẫu thể hiện bằng hai véc tơ trực giao trong khi thực tế cần phải lưu trữ, xử lý với cả các mẫu được thể hiện bằng hai véc tơ không trực giao.

3.3 Lý do đề xuất thuật toán học mới

Các nghiên cứu về BAM chia thành hai nhóm chính gồm nhóm các mô hình học một lần và nhóm mô hình học nhiều lần. Ở nhóm mô hình học một lần, do quá trình học không thực hiện dựa trên điều kiện về hàm năng lượng của BAM để đảm bảo khả năng nhớ lại đúng cặp mẫu đã lưu nên khả năng nhớ lại đúng (chịu nhiều) giảm rất nhiều đặc biệt khi mẫu vào có độ nhiễu cao hay mất nhiều thông tin.

Ở nhóm mô hình học nhiều lần, quá trình học chủ yếu dựa vào các điều kiện đảm bảo khả năng nhớ lại đúng cặp mẫu đã lưu. Nghĩa là, hàm năng lượng phải đạt cực tiểu cục bộ tại mỗi trạng thái ứng với mỗi cặp mẫu. Do đó, khả năng nhớ lại đúng các cặp mẫu đã lưu của các mô hình này khá cao ngay cả khi mẫu vào có mức độ nhiễu cao. Tuy nhiên, độ phức tạp tính toán của quá trình học lại khá lớn do tập mẫu được học đi học lại nhiều lần để nhớ tốt hơn các cặp mẫu.

Hơn nữa, các mô hình BAM trước đây nhớ lại tốt hơn với các cặp mẫu được thể hiện thành hai véc tơ trực giao. Do đó, khả năng nhớ lại đối với các cặp mẫu không trực giao là chưa được quan tâm.

Từ việc tổng hợp các nghiên cứu về BAM, tác giả nhận thấy BAM học nhiều lần sẽ phát huy tốt ưu điểm về phục hồi mẫu đã lưu từ mẫu vào nhiễu. Tuy nhiên, cần phải giảm thời gian học các mẫu trong khi vẫn giữ khả năng nhớ lại để đáp ứng cho các ứng dụng thời gian thực đặc biệt khi số lượng mẫu và kích thước của các mẫu tăng lên. Ngoài ra, BAM cần được cải tiến khả năng lưu trữ và nhớ lại các cặp mẫu được thể hiện thành hai véc tơ không trực giao.

Từ các phân tích trên, tác giả đưa ra một thuật toán học thể hiện chiến lược học nhiều lần với hai ưu điểm sau:

- Quá trình học các cặp mẫu thực hiện nhanh và linh hoạt hơn
- Mô hình BAM gắn với thuật toán học mới có khả năng nhớ lại tốt hơn với các cặp mẫu có thể hiện thành hai véc tơ không trực giao.

3.4 Thuật toán học mới cho BAM

3.4.1 Ý tưởng

Mục đích: Thuật toán học mới thực hiện học các cặp mẫu nhanh, linh hoạt. Mô hình BAM với thuật toán học mới (FFBAM - *Fast Flexible Bidirectional Associative Memory*) cải thiện khả năng nhớ lại các cặp mẫu không trực giao tốt hơn.

Ý tưởng: Thuật toán học mới thể hiện việc học nhiều lần các cặp mẫu thông qua số lần tối thiểu cần học một cặp mẫu (MNTP). Thuật toán học được thực hiện

như sau: Trong mỗi lần lặp của quá trình học, giảm MNTP của các cặp mẫu huấn luyện chưa đảm bảo điều kiện để BAM nhớ lại cặp mẫu huấn luyện.

Trong phần 2.1.2, công thức (2.28) và (2.29) thể hiện các trọng số liên kết trong \mathbf{W} và hàm năng lượng E bị ảnh hưởng bởi MNTP. Vì vậy, tác giả phân tích mối quan hệ giữa MNTP và hàm năng lượng thông qua ma trận trọng số liên kết \mathbf{W} .

3.4.2 Phân tích mối quan hệ giữa MNTP và hàm năng lượng

Giả sử, BAM học p cặp mẫu. Cặp mẫu $(\mathbf{A}^i, \mathbf{B}^i)$ được trình bày như sau: $\mathbf{A}^i = (A_1^i, \dots, A_n^i)$ và $\mathbf{B}^i = (B_1^i, \dots, B_m^i)$. Mối quan hệ giữa MNTP và hàm năng lượng được thiết lập từ công thức (2.28) và (2.29) có dạng như sau:

Trong công thức (2.29), $\mathbf{A}^k (\mathbf{B}^k)^T$ được thể hiện trong ma trận sau:

$$\mathbf{A}^k (\mathbf{B}^k)^T = \begin{bmatrix} A_1^k (B_1^k)^T & \dots & A_1^k (B_m^k)^T \\ \dots & \dots & \dots \\ A_n^k (B_1^k)^T & \dots & A_n^k (B_m^k)^T \end{bmatrix} \quad (3.9)$$

tiếp theo, ta có

$$q^k \mathbf{A}^k (\mathbf{B}^k)^T = \begin{bmatrix} q^k A_1^k (B_1^k)^T & \dots & q^k A_1^k (B_m^k)^T \\ \dots & \dots & \dots \\ q^k A_n^k (B_1^k)^T & \dots & q^k A_n^k (B_m^k)^T \end{bmatrix} \quad (3.10)$$

Thay vào công thức (2.29), ta có công thức tính \mathbf{W} :

$$\mathbf{W} = \begin{bmatrix} \sum_{k=1}^p q^k A_1^k (B_1^k)^T & \dots & \sum_{k=1}^p q^k A_1^k (B_m^k)^T \\ \dots & \dots & \dots \\ \sum_{k=1}^p q^k A_n^k (B_1^k)^T & \dots & \sum_{k=1}^p q^k A_n^k (B_m^k)^T \end{bmatrix} \quad (3.11)$$

Trong công thức (2.28), $\mathbf{A}^i \mathbf{W}$ được tính như sau:

$\mathbf{A}^i \mathbf{W}$

$$= \left(\sum_{t=1}^n A_t^i \sum_{k=1}^p q^k A_t^k (B_1^k)^T, \sum_{t=1}^n A_t^i \sum_{k=1}^p q^k A_t^k (B_2^k)^T, \dots, \sum_{t=1}^n A_t^i \sum_{k=1}^p q^k A_t^k (B_m^k)^T \right) \quad (3.12)$$

Sau đó, E^i được tính bằng công thức dưới đây:

$$\begin{aligned} E^i &= -\mathbf{A}^i \mathbf{W} \mathbf{B}^i{}^T \\ &= - \left(B_1^i \sum_{t=1}^n A_t^i \sum_{k=1}^p q^k A_t^k (B_1^k)^T + B_2^i \sum_{t=1}^n A_t^i \sum_{k=1}^p q^k A_t^k (B_2^k)^T \right. \\ &\quad \left. + \dots + B_m^i \sum_{t=1}^n A_t^i \sum_{k=1}^p q^k A_t^k (B_m^k)^T \right) \end{aligned}$$

Thu gọn công thức, ta có:

$$E^i = - \sum_{s=1}^m B_s^i \sum_{t=1}^n A_t^i \sum_{k=1}^p q^k A_t^k (B_s^k)^T \quad (3.13)$$

Công thức (3.13) thể hiện giá trị tuyệt đối của E^i tỷ lệ thuận với mọi q^k của p cặp mẫu (với $k=1, \dots, p$). Từ đó, suy ra được giá trị tuyệt đối của E^i sẽ giảm mạnh khi các q^k cùng giảm.

3.4.3 Nội dung thuật toán học mới

Theo kết quả nghiên cứu của Kosko, nếu năng lượng ở một trạng thái đạt đến một cực tiểu cục bộ thì BAM sẽ nhớ lại đúng cặp mẫu ứng với trạng thái đó. BAM được hình thành bởi việc ghép hai mạng Hopfield theo hai hướng ngược nhau. Theo hai định lý ở trang 33 và 34, hàm năng lượng của mạng Hopfield luôn giảm theo thời và giới nội nên năng lượng của mạng sẽ giảm đến một mức nhất định trong quá trình cập nhật.

Trong thuật toán học mới này, BAM sẽ học tất cả các cặp mẫu cho đến khi năng lượng của BAM ứng với mọi cặp mẫu đủ nhỏ. Do đó, khả năng nhớ lại mọi cặp mẫu là tương đương nhau (không phân biệt cặp mẫu được thể hiện thành hai véc tơ trực giao hay không) .

Một số ký hiệu trong thuật toán

- q^i là MNTP của cặp mẫu thứ i .
- \mathbf{W} là ma trận trọng số chung
- E^i là năng lượng ứng với trạng thái $(\mathbf{A}^i, \mathbf{B}^i)$.
- ε là ngưỡng để dừng việc điều chỉnh q^i .

Thuật toán 3.1: Thuật toán học nhanh và linh động cho BAM

Input: p cặp mẫu $(\mathbf{A}^i, \mathbf{B}^i)$ được thể hiện thành hai véc tơ chứa các giá trị ở dạng hai cực gồm +1 và -1.

Output: Ma trận trọng số \mathbf{W} lưu sự liên kết của các cặp mẫu

Nội dung thuật toán :

- **Bước 1:** Khởi tạo giá trị MNTP bằng cách đặt mọi $q^i=1$ với $i=1, \dots, p$ để thu được ma trận trọng số gốc. Chọn giá trị cho ε nguyên dương, đủ nhỏ.
- **Bước 2:** Thực hiện lặp các bước sau cho đến khi $|E^i| \leq \varepsilon$ với mọi $i=1, \dots, p$ và $|x|$ là giá trị tuyệt đối của x .

Bước 2.1: Tính \mathbf{W} theo công thức (3.11)

Bước 2.2: Tính E^i theo công thức (3.13) với $i=1, \dots, p$.

Bước 2.3: Dựa vào giá trị của E^i để cập nhật q^i theo hai luật sau:

Luật 1: Nếu $|E^i| \leq \varepsilon$ thì không thay đổi q^i

Luật 2: Nếu $|E^i| > \varepsilon$ thì giảm q^i xuống h lần với h là phần nguyên của phép chia $|E^i|$ cho ε .

- **Bước 3:** Trả về ma trận trọng số \mathbf{W}

Ưu điểm của thuật toán học mới

Ba lý do chính làm cho thuật toán học mới tốt hơn so với các thuật toán học đã công bố gồm: Sự tăng hay giảm trọng số kết nối được thực hiện bằng phép nhân hoặc phép chia; MNTP của mỗi cặp có thể được thay đổi trong mỗi lần lặp của quá

trình học; và năng lượng của mỗi trạng thái ứng với mỗi cặp mẫu là xấp xỉ của cực tiểu cục bộ. Vì vậy, FFBAM sẽ có hai ưu điểm tốt hơn các BAM khác.

Ưu điểm thứ nhất là *quá trình học các cặp mẫu thực hiện nhanh và linh hoạt hơn*. Quá trình học diễn ra nhanh trong một số rất ít các lần lặp do sự tăng hay giảm trọng số kết nối được thực hiện bằng phép nhân hoặc phép chia. Do đó, FFBAM nhanh chóng đạt điều kiện dừng của quá trình học (mọi $|E^i| \leq \epsilon$ với $i=1, \dots, p$.) khi giá trị MNTP của nhiều cặp mẫu cùng giảm. Hơn nữa, quá trình học là linh động do MNTP của mỗi cặp có thể được thay đổi trong mỗi lần lặp của quá trình học.

Trong các nghiên cứu trước đây [67,72,76], sự tăng hay giảm trọng số kết nối được thực hiện bằng phép cộng hoặc phép trừ và các tham số của mô hình trong quá trình học là không đổi. Vì vậy, thuật toán học đề xuất là khác hoàn toàn so với các thuật toán học nhiều lần đã có.

Ưu điểm thứ hai là *khả năng nhớ lại tốt hơn* khi xử lý với các cặp mẫu không thể hiện bằng hai véc tơ trực giao. Khả năng nhớ lại các cặp mẫu của FFBAM tốt hơn là do mọi cặp mẫu bất kỳ được học đều có giá trị năng lượng xấp xỉ đến giá trị cực tiểu toàn cục. Nghĩa là, điều kiện nhớ lại đúng về năng lượng là được thỏa mãn (ở mức gần đúng).

Theo kết quả của Kosko [45], năng lượng chỉ cần đạt đến cực tiểu cục bộ là có khả năng nhớ lại đúng các cặp mẫu được thể hiện thành hai véc tơ trực giao. Do đó, nếu BAM chỉ học và lưu trữ các cặp mẫu có thể hiện thành hai véc tơ trực giao thì các BAM khác có khả năng nhớ tốt hơn FFBAM. Tuy nhiên, nếu các mẫu học có cả các mẫu không có thể hiện thành hai véc tơ trực giao thì FFBAM nhớ tốt hơn do mọi cặp mẫu đều có khả năng nhớ lại giống nhau.

Trong các nghiên cứu trước đây, quá trình học dừng khi năng lượng của BAM ứng với mỗi trạng thái phải đạt đến cực tiểu cục bộ [68,69,70] hoặc dừng khi các trọng số kết nối ổn định [67,76]. Do đó, điều kiện dừng quá trình học của tác giả là khác hoàn toàn so với các nghiên cứu đã có.

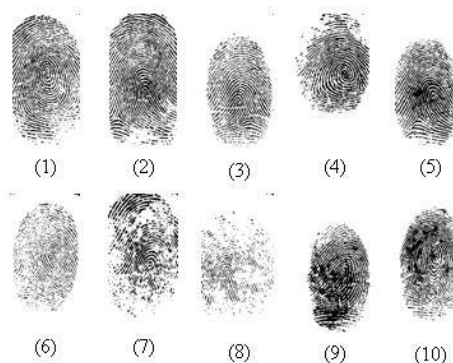
3.5 Kết quả thực nghiệm

Tác giả thử nghiệm các BAM với ứng dụng nhận dạng. Các ảnh mẫu gốc và các ảnh nhiễu vào được chuyển đổi thành ảnh đen trắng, sau đó chuyển thành các véc tơ một chiều với các giá trị là 1 hoặc -1.

Do mô hình BAM gắn với thuật toán học mới thực hiện học nhiều lần các mẫu nên tác giả chỉ so sánh với các BAM học nhiều lần. Các BAM được dùng trong các thực nghiệm gồm BAM của Y.F.Wang (WBAM) [69], BAM của Zhuang (ZBAM) [76], BAM của Tao Wang (TBAM) [67] và FFBAM. Các thử nghiệm ở chế độ tự liên kết được làm để phù hợp với ứng dụng nhận dạng thực tế. Để đo thời gian học các mẫu, tác giả dùng công cụ đo trong Matlab để tính thời gian tính toán đưa ra được ma trận trọng số W . Đối với khả năng nhớ lại, tác giả sử dụng tỷ lệ phần trăm số điểm ảnh được nhớ lại đúng.

3.5.1 Thử nghiệm với nhận dạng vân tay

Tập mẫu gồm 10 ảnh huấn luyện được chọn từ tập mẫu của cuộc thi Olympic về công nghệ thông tin tổ chức tại ĐH Công nghệ năm 2010. Mỗi ảnh được giảm kích thước về còn 23x12. 10 ảnh nhiễu được tạo ra từ mỗi ảnh huấn luyện bằng việc trao đổi giá trị của một số điểm ảnh với mức độ tăng dần. Hình 3.1 thể hiện các ảnh huấn luyện và Bảng 3.1 thể hiện kết quả thử nghiệm về thời gian học và khả năng nhớ lại.



Hình 3.1: Các ảnh huấn luyện trong nhận dạng vân tay.

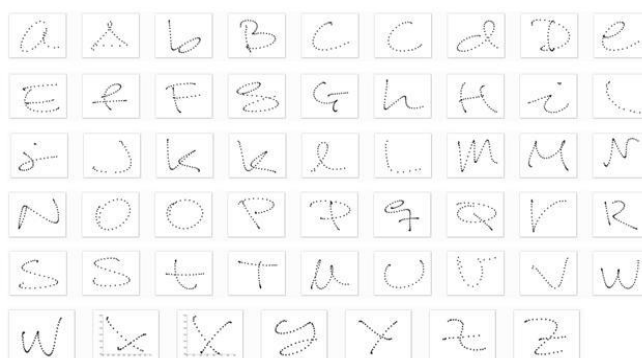
Bảng 3.1: Thời gian học và kết quả nhớ lại các vân tay

Mô hình được so sánh	WBAM	ZBAM	TBAM	FFBAM
Tỷ lệ nhớ lại đúng (%)	83.370	85.906	85.906	88.007
Thời gian học (s)	0.054	6.927	161.164	0.648

Bảng 3.1 cho thấy FFBAM là mô hình tốt nhất. Khả năng nhớ của FFBAM cao hơn WBAM khoảng 4.6% nhưng học chậm hơn khoảng 0.6s. Đối với hai mô hình còn lại, việc học thực hiện nhanh hơn khoảng 6s và nhớ lại tốt hơn khoảng 2%.

3.5.2 Thử nghiệm với nhận dạng chữ viết tay

Tập mẫu gồm 52 ảnh được chọn từ tập mẫu UJIpenchars từ cơ sở dữ liệu UCI trong đó có 26 chữ thường và 26 chữ hoa. Mỗi ảnh được giảm kích thước về còn 13x17. Tác giả sử dụng chữ viết tay của 4 người khác để làm ảnh nhiễu. Hình 3.2 thể hiện các ảnh huấn luyện và Bảng 3.2 thể hiện thời gian học và khả năng nhớ lại của các BAM. Tương tự, Bảng 3.2 cho thấy khả năng nhớ của FFBAM là cao nhất và chỉ học chậm hơn WBAM 0.02s.



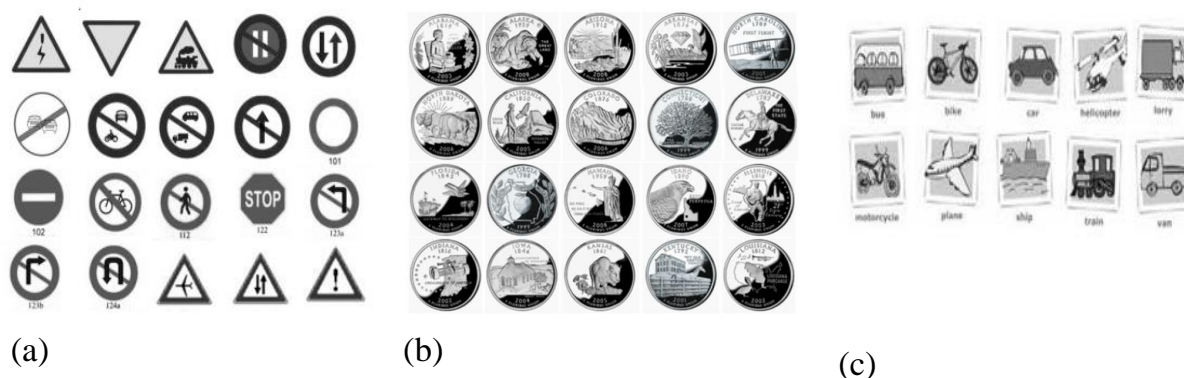
Hình 3.2: Các ảnh huấn luyện trong nhận dạng chữ viết tay.

Bảng 3.2: Thời gian học và kết quả nhớ lại các chữ viết tay

Mô hình được so sánh	WBAM	ZBAM	TBAM	FFBAM
Tỷ lệ nhớ lại đúng (%)	75.463	72.964	75.681	75.89
Thời gian học (s)	0.195	153.546	198.955	0.212

3.5.3 Thử nghiệm với các ứng dụng nhận dạng khác

Tập mẫu gồm 20 ảnh về biển hiệu giao thông, 20 ảnh về tiền xu của Mỹ và 10 ảnh về phương tiện giao thông được chọn từ kết quả tìm kiếm ảnh trên Google. Mỗi ảnh huấn luyện được giảm kích thước về còn 23x23. 10 ảnh nhiễu được tạo ra từ mỗi ảnh huấn luyện bằng cách trao đổi giá trị điểm ảnh với số điểm bị thay đổi là ngẫu nhiên. Hình 3.3 thể hiện các ảnh huấn luyện và Bảng 3.3, 3.4, và 3.5 thể hiện khả năng nhớ lại của các BAM.



Hình 3.3: Các ảnh huấn luyện trong các ứng dụng nhận dạng khác. Hình 3.3(a) – Các ảnh mẫu về biển hiệu giao thông, Hình 3.3(b) – Các ảnh mẫu về tiền xu của Mỹ, và Hình 3.3(c) – Các ảnh mẫu về các phương tiện giao thông

Bảng 3.3: Thời gian học và kết quả nhớ lại các biển hiệu giao thông

Mô hình được so sánh	WBAM	ZBAM	TBAM	FFBAM
Tỷ lệ nhớ lại đúng (%)	77.98	28.303	78.303	78.348
Thời gian học (s)	0.057	59.751	17.020	0.409

Bảng 3.4: Thời gian học và kết quả nhớ lại các tiền xu của Mỹ

Mô hình được so sánh	WBAM	ZBAM	TBAM	FFBAM
Tỷ lệ nhớ lại đúng (%)	85.066	45.992	84.896	85.109
Thời gian học (s)	0.332	55.291	100.520	0.815

Bảng 3.5: Thời gian học và kết quả nhớ lại các phương tiện giao thông

Mô hình được so sánh	WBAM	ZBAM	TBAM	FFBAM
Tỷ lệ nhớ lại đúng (%)	88.11	18.96	90.076	90.076
Thời gian học (s)	0.134	55.755	6.008	3.414

Tương tự, Bảng 3.3, 3.4, và 3.5 cho thấy khả năng nhớ của FFBAM là cao nhất và học chậm hơn WBAM trong cả ba thử nghiệm.

Trong các thực nghiệm, tác giả đã cố gắng chọn nhiều tập ảnh huấn luyện khác nhau để đa dạng hóa nội dung các mẫu lưu trữ trong BAM và chọn những ứng dụng có khả năng áp dụng BAM tốt nhất. Kết quả thực nghiệm cho thấy FFBAM có khả năng nhớ lại tốt hơn các mô hình khác trong cả 5 thực nghiệm và chỉ học chậm hơn một mô hình (trường hợp chậm nhất là 3.3s). Điều này cho thấy FFBAM cải tiến được chất lượng xử lý của các ứng dụng nhận dạng với các mẫu đơn giản.

2.6 Kết luận chương

Hiện nay, các ứng dụng nhận dạng ngày càng được nâng cao về mặt chất lượng nên hệ thống có thể nhận dạng được các đối tượng từ các mẫu vào phức tạp như nhận dạng hành động của nhân vật trong các đoạn phim hay nhận dạng sự thay đổi thời tiết thông qua các ảnh vệ tinh,... BAM là một ANN đáp ứng được các yêu cầu này khi chúng ta giảm bớt chất lượng ảnh về dạng ảnh đen trắng. Tuy nhiên để nhận dạng tốt các mẫu từ các mẫu vào nhiễu, các BAM đã công bố đòi hỏi thời gian học các mẫu lớn. Do đó, việc phát triển một thuật toán học nhanh hơn trong khi vẫn giữ được khả năng nhớ lại các mẫu là hết sức cần thiết cho các ứng dụng nhận dạng mẫu.

Tác giả đã đưa ra một thuật toán thực hiện học nhanh và linh động hơn. Khả năng nhớ lại của BAM gắn với thuật toán học mới cao hơn các BAM học nhiều lần đã công bố trong chế độ tự liên kết, đặc biệt khi BAM lưu các cặp mẫu được thể hiện bằng hai véc tơ không trực giao. Kết quả nghiên cứu này được công bố tại kỷ yếu có phản biện của Hội nghị quốc tế lần thứ 8 về Machine Learning và Data

Mining - MLDM 2012 (Công trình khoa học số 2), kỷ yếu có phản biện của Hội nghị quốc tế lần thứ nhất về Information Technology and Science (Công trình khoa học số 1) , và Tạp chí Khoa học Công nghệ trong nước (Công trình khoa học số 7).

CHƯƠNG 4. HAI LUẬT HỌC CẢI TIẾN CHO LÝ THUYẾT CỘNG HƯỞNG THÍCH NGHI MỜ

Trong phần 2.6 và 2.7 của chương 2, tác giả đã trình bày các hiểu biết quan trọng về ART và Fuzzy ART. Trong chương này, tác giả sẽ trình bày các nghiên cứu có liên quan đến ART và Fuzzy ART để làm cơ sở lý luận đề xuất hai luật học cho Fuzzy ART. Tiếp theo, nội dung của hai luật học cải tiến và thuật toán tìm giá trị thích hợp cho tham số học của Fuzzy ART được thể hiện trong phần 4.4. Phần 4.5 sẽ mô tả và phân tích chi tiết các kết quả thực nghiệm.

4.1 Giới thiệu chung

ART là một ANN học các mẫu huấn luyện để hình thành véc tơ trọng số của các cụm. Sau đó, sử dụng véc tơ trọng số của các cụm để nhớ lại các cụm tương tự với mẫu vào. Ưu điểm quan trọng nhất của ART là học mẫu huấn luyện để véc tơ trọng số của cụm chiến thắng chứa các thông tin mới từ mẫu huấn luyện hiện tại. Nghĩa là, trọng số của các cụm được cập nhật cho đến khi trọng số của các cụm gần hơn với mẫu huấn luyện. Fuzzy ART là một dạng ART có sử dụng phép giao của logic mờ và các mẫu được thể hiện ở dạng tập mờ. Do đó, Fuzzy ART vẫn có được ưu điểm của ART và tận dụng được các kết quả nghiên cứu về mặt toán học của logic mờ để chọn được giá trị thích hợp cho các tham số của mô hình.

4.2 Các nghiên cứu liên quan

4.2.1 Mô hình ART

Các nghiên cứu về ART được phát triển để cải thiện khả năng phân cụm. Tan [61] thể hiện một kiến trúc nơ-ron được gọi là ánh xạ liên kết cộng hưởng thích nghi (ARAM). Kiến trúc này mở rộng các hệ thống ART không giám sát về tốc độ, tính ổn định và học các liên kết khác loại. Với việc thiết lập ngưỡng cực đại, ARAM mã hóa các cặp mẫu chính xác và nhớ lại hoàn hảo. Hơn nữa, mô hình này tạo ra tính chịu nhiễu mạnh. Kenaya và Cheok [41] đưa ra mô hình Euclidean ART trong đó dùng khoảng cách Euclid để đo sự tương tự một mẫu với véc tơ trọng số của các cụm đang có. Sau đó, quyết định mẫu thuộc về một cụm đang tồn tại hay

hình thành một cụm mới. Lin và đồng nghiệp [48] đưa ra thuật toán học dựa vào ART để phân chia trực tuyến không gian tín hiệu vào-tín hiệu ra của một bộ điều khiển mờ truyền thống. Mô hình này thiết lập hàm thành viên và tìm ra các luật logic mờ thích hợp dựa vào sự phân bố dữ liệu.

4.2.2 Mô hình Fuzzy ART

Các nghiên cứu về lý thuyết của Fuzzy ART có thể chia thành ba nhóm, gồm phát triển các mô hình lý thuyết, nghiên cứu các thuộc tính, và tối ưu hóa sự thực hiện.

Các mô hình lý thuyết của Fuzzy ART

Các mô hình Fuzzy ART được đưa ra để cải thiện khả năng phân cụm dữ liệu. Carpenter và đồng nghiệp [8] đưa ra Fuzzy ARTMAP cho việc học các cụm nhận dạng và các ánh xạ đa chiều các mẫu vào nhị phân hay tương tự. Mô hình này cực tiểu lỗi dự đoán và cực đại sự tổng quát hóa các mẫu. Do đó, hệ thống tự học số cụm ít nhất đạt yêu cầu về độ chính xác. Việc dự đoán được cải thiện do việc huấn luyện hệ thống nhiều lần với nhiều thứ tự khác nhau của tập dữ liệu vào, sau đó chọn một. Chiến lược chọn này có thể dùng để ước lượng dự đoán với các tập mẫu huấn luyện không đầy đủ hay các tập mẫu nhỏ và nhiễu. Lin và đồng nghiệp [48] đưa ra thuật toán học theo cấu trúc hoặc học theo tham số để xây dựng ANN truyền thẳng nhiều tầng cho nhận dạng các thành phần của bộ điều khiển. Do đó, nhóm tác giả đưa ra thuật toán học dựa vào Fuzzy ART để chia không gian dữ liệu vào-dữ liệu ra một cách linh động dựa vào sự phân bố của dữ liệu. Thuật toán này phối hợp học theo tham số của mạng lan truyền ngược và học cấu trúc của thuật toán Fuzzy ART. Hơn nữa, thuật toán này điều chỉnh các hàm thành viên, và tìm ra các luật logic mờ thích hợp. Isawa và đồng nghiệp [38] đưa ra một bước thêm, học nhóm, cho Fuzzy ART để thu được kết quả phân cụm hiệu quả hơn. Đặc trưng quan trọng của học nhóm là tạo ra các kết nối giữa các cụm tương tự. Nghĩa là mô hình này học cả các cụm và sự kết nối giữa các cụm. Sau đó, nhóm tác giả này [39] đưa ra một Fuzzy ART phối hợp các cụm xếp chồng trong các kết nối để tránh vấn đề sinh cụm mới. Đặc trưng quan trọng của nghiên cứu này là sắp xếp các tham số ngưỡng

cho mọi cụm và thay đổi các tham số ngưỡng này theo kích thước của các cụm trong quá trình học. Yousuf and Murphey [73] đưa ra một thuật toán so sánh các trọng số của các cụm với mẫu vào và cho phép cập nhật nhiều cụm thỏa mãn điều kiện về ngưỡng. Mô hình này vượt qua các hạn chế của Fuzzy ART gốc do một mẫu có thể thuộc về nhiều cụm. Hai tác giả mô phỏng sự ảnh hưởng của việc cập nhật trọng số của các cụm khi phân cụm sai và đề xuất việc phạt khi cập nhật trọng số sai. K.L. Chu và đồng nghiệp [12] đã cải tiến Fuzzy ARTMAP để xử lý với các giá trị phức tạp của dữ liệu không gian-thời gian thu được từ mắt người. Mô hình cho phép bảo toàn các mẫu đã học trong khi học các mẫu mới. Các thực nghiệm được làm để nhận dạng đối tượng từ các ảnh đã được căn chỉnh và các ảnh ko căn chỉnh. V.Vidya và đồng nghiệp [63] thực hiện nhận dạng chữ viết tay của ngôn ngữ Malayalam bằng cách dùng kết hợp một mạng Fuzzy ARTMAP và mạng nơ-ron Bayer. Hơn nữa, tối ưu hóa bầy đàn từng phần được áp dụng để cải thiện khả năng phân lớp ký tự chính xác hơn. W.Y. Shen và đồng nghiệp [55] đã phối hợp Fuzzy ART và máy học trực tuyến để cung cấp khả năng tổng quát hóa lớn hơn và học nhanh hơn trong nhiệm vụ phân lớp mẫu. Mô hình này tự động sinh các nơ-ron ẩn để chứa các thông tin mới mà không xếp chồng và lẫn với các tri thức đã học. K.L. Chu và đồng nghiệp [13] đề xuất một phương thức phân lớp tối ưu bằng việc giải quyết vấn đề thiết kế của Fuzzy ART về thứ tự của các mẫu huấn luyện và tối ưu tham số của các ngưỡng trong mô hình. Thuật toán di truyền và thông tin xác xuất được dùng để cải thiện khả năng phân lớp.

Trong hướng nghiên cứu này, luật học trong các mô hình còn học chưa hiệu quả do các mẫu huấn luyện có giá trị của các phần tử lớn hơn giá trị của các phần tử tương ứng trong các véc tơ trọng số của các cụm không có ảnh hưởng đến cụm. Lý do là trong luật học dùng phép toán giao của logic mờ giữa véc tơ thể hiện mẫu vào và véc tơ trọng số của cụm được chọn. Hệ quả là khả năng phân cụm của Fuzzy ART sẽ giảm khi nhiều mẫu huấn luyện quan trọng hoặc có ý nghĩa không được lưu trữ trong quá trình học

Các thuộc tính của Fuzzy ART

Các thuộc tính quan trọng được nghiên cứu để chọn ra các tham số phù hợp cho một Fuzzy ART mới. Huang và đồng nghiệp [37] biểu diễn một số thuộc tính quan trọng của Fuzzy ART để cung cấp nhiều hiểu biết về các thao tác của Fuzzy ART. Hơn nữa, các ảnh hưởng của tham số chọn và tham số ngưỡng đến chức năng của thuật toán Fuzzy ART được trình bày rõ hơn. Georgiopoulos và đồng nghiệp [23] tập trung vào thứ tự các cụm sẽ được chọn. Nghiên cứu này cung cấp lý do tại sao và thứ tự nào các cụm được chọn tùy vào vùng giá trị của tham số chọn. Anagnostopoulos và Georgiopoulos [3] giới thiệu các khái niệm hình học (các vùng thể hiện mỗi cụm) trong khung công việc gốc của Fuzzy ART và Fuzzy ARTMAP. Sau đó, định nghĩa các vùng này dựa vào các biểu diễn hình học của các điều kiện về ngưỡng và sự cạnh tranh của các nút được chọn với các nút chưa được chọn. Kết quả thực tế là một trong các trạng thái của không gian tham số ngưỡng-tham số chọn cho kết quả của quá trình học và nhớ lại không phụ thuộc vào các lựa chọn cụ thể của tham số ngưỡng.

Tối ưu sự thực hiện của Fuzzy ART

Trong nhóm nghiên cứu này, các nghiên cứu tập trung vào cải thiện sự thực hiện của Fuzzy ART. Burwick và Joublin [6] thảo luận các thực thi của ART trên một máy nối tiếp. Tính toán chuẩn của ART tương ứng với một thuật toán đệ quy với độ phức tạp $O(N*N+M*N)$. Nghiên cứu này làm cho độ phức tạp thuật toán giảm xuống còn $O(N*M)$ thông qua một thuật toán không đệ quy với N là số cụm và M là số chiều của mẫu vào. Dagher và đồng nghiệp [17] giới thiệu một thủ tục dựa trên phương thức phân cụm max-min cho Fuzzy ARTMAP. Mô hình này có sự tổng quát hóa tốt nhưng số lượng thao tác lại rất nhỏ so với Fuzzy ARTMAP. Cano và đồng nghiệp [7] sinh ra các bộ xác định hàm chính xác cho các dữ liệu nhiễu dựa vào ART. Sau đó, một ANN truyền thẳng và một ART được đưa ra. Hai mô hình này có thể huấn luyện trên dữ liệu nhiễu trong khi không cần thay đổi cấu trúc và không cần tiền xử lý dữ liệu. Kobayashi [42] đưa ra một hệ thống học tăng cường

mới có sử dụng Fuzzy ART để phân lớp các thông tin quan sát và xây dựng một không gian trạng thái hiệu quả.

4.2.3 Các luật học điển hình của ART và Fuzzy ART

Các mô hình sử dụng một số biến sau: \mathbf{I} là mẫu vào hiện tại, \mathbf{W}_j là trọng số của cụm j , và β là tham số học nằm trong khoảng $[0, 1]$.

Capenter và đồng nghiệp [9] đưa ra mô hình Fuzzy ART gốc với luật học cập nhật trọng số cho cụm được chọn j như sau:

$$\mathbf{W}_j^{\text{new}} = \beta(\mathbf{I} \wedge \mathbf{W}_j^{\text{old}}) + (1 - \beta)\mathbf{W}_j^{\text{old}} \quad (4.1)$$

với \wedge là phép giao của hai tập mờ.

Hầu hết các nghiên cứu khác cũng dùng luật học như công thức (4.1) như Capenter và đồng nghiệp [8], A.H.Tan [61], Isawa và đồng nghiệp [38]. Ngoài ra, các nghiên cứu này tập trung vào thay đổi hàm đầu ra hay học bổ sung thêm quan hệ giữa các nhóm các cụm

Kenaya và Cheok [41] đưa ra Euclidean ART để học với các dữ liệu nhiễu với luật học như sau:

$$\mathbf{W}_j^{\text{new}} = \frac{\sum_{k=1}^L X_{jk}}{L} \quad (4.2)$$

với X_{jk} là mẫu thứ k trong cụm j và L là số các cụm.

Yousuf and Murphey [73] cập nhật nhiều cụm thỏa mãn điều kiện về ngưỡng theo luật học sau:

$$\mathbf{W}_j^{\text{new}} = \beta(\mathbf{Y}_j(\mathbf{I}) \wedge \mathbf{W}_j^{\text{old}}) + (1 - \beta)\mathbf{W}_j^{\text{old}} \quad (4.3)$$

với $\mathbf{Y}_j(\mathbf{I}) = \frac{\|\mathbf{I} \wedge \mathbf{W}_j\|}{\alpha + \|\mathbf{W}_j\|}$

Trong các nghiên cứu trên, trọng số mới của các cụm bằng tỷ lệ phần trăm của trọng số cũ của các cụm cộng với tỷ lệ phần trăm còn lại của giá trị cực tiểu giữa véc tơ thể hiện mẫu vào và trọng số cũ của các cụm. Do đó, hai luật học trong (4.1) và (4.3) chưa thể hiện rõ ảnh hưởng của các mẫu huấn luyện đến trọng số của các cụm. Đặc biệt là khi mọi giá trị trong véc tơ biểu diễn mẫu vào là lớn hơn các giá trị

tương ứng trong véc tơ trọng số của các cụm thì trọng số của các cụm không thay đổi. Nghĩa là, thông tin về mẫu huấn luyện hiện tại không được lưu lại trong véc tơ trọng số cụm. Do đó, khi số mẫu không được học nhiều hoặc một số mẫu quan trọng không được học sẽ làm giảm khả năng phân cụm của Fuzzy ART.

4.3 Lý do đề xuất hai luật học

Dựa vào sự tổng hợp các nghiên cứu về ART, và Fuzzy ART, tác giả nhận thấy có hạn chế sau: Luật học của ART và Fuzzy ART là chưa hiệu quả do một số mẫu huấn luyện không có ảnh hưởng đến trọng số cụm. Do đó, nhiều thông tin có ích từ các mẫu huấn luyện đó đã bị mất mát.

Từ các phân tích trên, tác giả đề xuất hai luật học cho Fuzzy ART để học mọi mẫu huấn luyện. Hơn nữa, các luật học này cũng giảm sự ảnh hưởng của các mẫu huấn luyện dị thường đến trọng số của các cụm. Để cho quá trình học được hiệu quả, một thuật toán tìm giá trị thích hợp cho tham số tốc độ học được trình bày.

4.4 Hai luật học đề xuất cho Fuzzy ART

4.4.1 Ý tưởng

Mục đích: đưa ra luật học cải tiến để lưu trữ mọi mẫu huấn luyện và giảm sự ảnh hưởng của các mẫu dị thường đối với véc tơ trọng số của các cụm. Fuzzy ART với luật học mới có thể cải thiện khả năng phân cụm.

Ý tưởng: Luật học đề xuất thể hiện trọng số của các cụm được tăng hoặc giảm để cụm chiến thắng thích nghi được với mẫu huấn luyện. Sự ảnh hưởng của mỗi mẫu huấn luyện đối với cụm chiến thắng được thể hiện bởi tham số học. Một thuật toán tìm ra một giá trị thích hợp cho tham số học đối với mỗi tập dữ liệu được đưa ra. Trong thuật toán này, ban đầu tham số học được khởi tạo dựa vào kích thước của tập dữ liệu để giảm thời gian tính toán. Sau đó, tham số học được tăng/giảm cho đến khi khả năng phân cụm đạt mức lỗi đủ nhỏ.

4.4.2 Nội dung của hai luật học

Sau khi Fuzzy ART chọn được một cụm chiến thắng, việc học mẫu huấn luyện hiện tại diễn ra. Giả sử, cụm chiến thắng là cụm j .

Luật học thứ nhất

Thực hiện cập nhật trọng số cho cụm j theo công thức dưới đây:

$$W_{ji}^{new} = W_{ji}^{old} - \delta |I_i - W_{ji}^{old}|, i = 1, \dots, M \quad (4.4)$$

với δ là tham số học tốc độ học và $|y|$ là giá trị tuyệt đối của y .

Sau khi cập nhật có thể điều chỉnh W_{ij} theo luật sau: Do W_{ij} luôn giảm nên khi $W_{ij} < 0$ thì đặt $W_{ij} = 0$.

Luật học thứ hai

Trước tiên, tính sự tăng giá trị bé nhất (MDI- *the Minimum Difference of Increase*) và sự giảm giá trị bé nhất (MDD - *the Minimum Difference of Decrease*) của mẫu vào hiện tại so với trọng số của cụm chiến thắng theo các công thức sau:

$$MDD = \min_{I_i \leq W_{ji}^{old}, i=1, \dots, M} W_{ji}^{old} - I_i \quad (4.5)$$

$$MDI = \min_{I_i > W_{ji}^{old}, i=1, \dots, M} I_i - W_{ji}^{old} \quad (4.6)$$

Khi đó, luật học thứ hai được trình bày như sau:

$$W_{ji}^{new} = \begin{cases} W_{ji}^{old} - \delta * MDD, I_i < W_{ji}^{old} \\ W_{ji}^{old}, I_i = W_{ji}^{old} \\ W_{ji}^{old} + \delta * MDI, I_i > W_{ji}^{old} \end{cases} \quad (4.7)$$

Thuật toán 4.1: Tìm giá trị thích hợp cho tham số tốc độ học của Fuzzy ART

Ý tưởng: Chọn ra một tập con ngẫu nhiên các mẫu. Cho Fuzzy ART thực hiện phân cụm với từng giá trị của tham số tốc độ học. Dùng chỉ số Davies-Bouldin trong công thức (2.53) để đo chất lượng phân cụm. Nếu chỉ số Davies-Bouldin nhỏ hơn một ngưỡng ε nguyên dương thì dừng việc tìm kiếm.

Input: Một tập mẫu con ngẫu nhiên từ tập dữ liệu ban đầu.

Output: Giá trị thích hợp của tham số tốc độ học.

Nội dung thuật toán

Bước 1: Khởi tạo giá trị cho các biến

Bước 1.1: Thiết lập giá trị cho tham số tốc độ học dựa vào kích thước của tập dữ liệu và miền giá trị của các phần tử thể hiện mẫu.

Bước 1.2: Thiết lập giá trị cho biến ε nguyên dương đủ nhỏ.

Bước 2: Lặp lại các bước sau:

Bước 2.1: Tính kết quả phân cụm của tập mẫu.

Bước 2.2: Kiểm tra chỉ số Davies–Bouldin về chất lượng phân cụm:

- Nếu chỉ số Davies–Bouldin index lớn hơn ε thì làm Bước 2.3.
- Nếu chỉ số Davies–Bouldin index nhỏ hơn ε thì dừng thuật toán và đưa ra giá trị của tham số tốc độ học

Bước 2.3: Thay đổi giá trị của tham số tốc độ học theo các bước nhảy nhỏ như sau: ở lần thay đổi đầu tiên, giảm giá trị của tham số. Với các lần thay đổi sau, có 4 trường hợp:

- Lần thay đổi trước thực hiện giảm giá trị và chất lượng phân cụm tăng thì tiếp tục giảm giá trị.
- Lần thay đổi trước thực hiện giảm giá trị và chất lượng phân cụm giảm thì tăng giá trị.
- Lần thay đổi trước thực hiện tăng giá trị và chất lượng phân cụm tăng thì tiếp tục tăng giá trị.
- Lần thay đổi trước thực hiện tăng giá trị và chất lượng phân cụm giảm thì giảm giá trị.

(Theo kinh nghiệm thu được từ các thực nghiệm nên chọn bước nhảy là 5% giá trị của tham số tốc độ học hiện tại).

4.4.3 Ưu điểm của hai luật học

Với hai luật học cải tiến này, mọi mẫu huấn luyện đều có một mức độ ảnh hưởng giống nhau đến trọng số của cụm được chọn do tham số học cố định trong suốt quá trình học. Do đó, mọi mẫu huấn luyện đều được lưu lại.

Với các mẫu dị thường có các giá trị thể hiện mẫu nhỏ hơn nhiều so với các giá trị thể hiện trọng số của cụm thì sự chênh lệch giá trị tương ứng của mẫu huấn luyện và trọng số cụm sẽ khá lớn. Vì vậy, khi dùng luật học gốc với phép giao hai tập mờ thì trọng số cụm mới sẽ bị kéo ra khá xa so với trọng số cụm cũ trong khi trọng số cụm cũ được hình thành do học nhiều mẫu huấn luyện trước đó.

Do việc các mẫu đều được lưu và chọn được tham số học thích hợp nên khả năng phân cụm của Fuzzy ART gắn với luật học đề xuất cũng sẽ được cải thiện.

4.5 Kết quả thực nghiệm

Các tập dữ liệu chuẩn từ cơ sở dữ liệu UCI¹ và Shape² được dùng trong các thử nghiệm để đánh giá tính hiệu quả của các mô hình được so sánh. Fuzzy ART của Carpenter được thực thi thành hai mô hình gồm mô hình thứ nhất là Original Fuzzy ART và mô hình thứ hai là Complement Original Fuzzy ART. Tương tự, mô hình Fuzzy ART với luật học đề xuất (EFART - Effective Fuzzy ART) cũng thực thi thành hai mô hình gồm Original EFART và Complement EFART. Tác giả sử dụng các mô hình sau trong các thực nghiệm: Original EFART (OriEFART), Complement EFART (ComEFART), Original Fuzzy ART (OriFART) [9], Complement Fuzzy ART (ComFART) [9], K-mean [51], và Euclidean ART (EucART) [41] để chứng minh tính hiệu quả của EFART. Giá trị của các tham số trong các mô hình so sánh được chọn để đạt được chất lượng phân cụm tốt nhất (giá trị chỉ số Davies-Bouldin nhỏ hơn một ngưỡng nhất định).

Dữ liệu của các tập dữ liệu được chuẩn hóa về miền $[0,1]$. Thực nghiệm áp dụng cho bài toán phân lớp các tập dữ liệu nên tập dữ liệu vào là các tập dữ liệu chuẩn đã chọn nhưng đã bỏ đi các nhãn lớp còn đầu ra là nhãn lớp của từng mẫu trong tập dữ liệu được kiểm tra.

Với mỗi tập dữ liệu, một mẫu ngẫu nhiên của mỗi phân lớp được chọn làm véc tơ trọng số ban đầu. Các thử nghiệm con được làm với số lượng mẫu tăng dần. Tỷ lệ

¹ Dữ liệu có tại địa chỉ <http://archive.ics.uci.edu/ml/datasets>

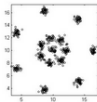
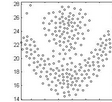
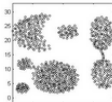
² Dữ liệu có tại địa chỉ <http://cs.joensuu.fi/sipu/datasets/>

phần trăm các mẫu được phân lớp đúng được thể hiện trong một bảng tương ứng với mỗi tập dữ liệu. Các số in đậm trong mỗi bảng thể hiện kết quả phân lớp của mô hình tốt nhất.

4.5.1 Thử nghiệm 1: Dùng luật học thứ nhất

9 tập dữ liệu chuẩn được chọn từ cơ sở dữ liệu UCI và Shape bao gồm Iris, Wine, Jain, Flame, R15, Glass, Blance-Scale, Aggregation, và Spiral. Các tập dữ liệu này khác nhau số thuộc tính, số lớp, số mẫu huấn luyện, và sự phân bố các mẫu ở các lớp. Bảng 4.1 thể hiện các thông tin trên của các tập dữ liệu được chọn.

Bảng 4.1: Đặc trưng của các tập dữ liệu trong thử nghiệm 1

Thứ tự	Tên tập dữ liệu	Số lớp	Số thuộc tính	Số mẫu	Nội dung dữ liệu
1	Iris	3	4	150	Hoa phong lan
2	Glass	7	9	214	Các loại kính bị vỡ để điều tra
3	Wine	3	13	178	Nguồn gốc rượu vang
4	Jain	2	2	373	Sự co cụm của người dùng mạng
5	Spiral	3	2	312	Phổ của các ảnh màu
6	Blance-Scale	3	4	625	Thử nghiệm tâm lý học
7	R15	15	2	600	 Các điểm dữ liệu tạo hình
8	Flame	2	2	240	 Các điểm dữ liệu tạo hình
9	Aggregation	7	2	788	 Các điểm dữ liệu tạo hình

Kiểm tra với tập Iris

Sự phân bố số mẫu trong ba lớp là đồng đều, mỗi lớp có 50 mẫu. Bảng 4.2 thể hiện kết quả thực nghiệm với tập mẫu Iris. Các kết quả ở Bảng 4.2 thể hiện rằng Complement EFART thực hiện tốt nhất trong mọi kiểm tra con.

Bảng 4.2: Kết quả phân lớp đúng của tập Iris

Số mẫu	OriEFART	ComEFART	OriFART	ComFART	EucART	K-mean
30	100	100	100	100	100	100
60	98.3	100	91.7	100	96.7	100
90	93.3	96.7	72.2	92.2	90.0	94.4
120	95.0	95.8	73.3	92.5	90.0	93.3
150	96.0	95.3	78.7	92.7	90.0	93.3

Kiểm tra với tập Spiral

Sự phân bố số mẫu trong ba lớp là khá đồng đều gồm 101, 105, và 106. Các số liệu từ Bảng 4.3 thể hiện rằng khả năng phân lớp đúng của Original EFART là tốt nhất trong mọi kiểm tra con, ngoại trừ kiểm tra con cuối cùng (Original EFART thấp hơn 4.1% so với mô hình tốt nhất - Euclidean ART).

Bảng 4.3: Kết quả phân lớp đúng của tập Spiral

Số mẫu	OriEFART	ComEFART	OriFART	ComFART	EucART	K-mean
50	88.0	4.0	4.0	22.0	2.0	44.0
100	71.0	49.0	37.0	25.0	48.0	22.0
150	48.0	33.3	25.3	21.3	32.7	16.7
200	42.5	42.0	29.5	39.0	41.0	32.0
250	40.4	33.6	27.6	35.2	39.6	32.0
312	38.5	37.8	33.0	28.2	42.6	32.7

Kiểm tra với tập Flame

Sự phân bố số mẫu trong hai lớp là 87 và 153. Bảng 4.4 thể hiện rằng Original EFART phân lớp tốt nhất trong mọi kiểm tra con, ngoại trừ kiểm tra con cuối cùng (Original EFART thấp hơn 3.3% so với mô hình tốt nhất - Original Fuzzy ART).

Bảng 4.4: Kết quả phân lớp đúng của tập Flame

Số mẫu	OriEFART	ComEFART	OriFART	ComFART	EucART	K-mean
50	100	100	100	76.0	88.0	78.0
100	98.0	87.0	87.0	83.0	94.0	54.0
150	98.7	87.3	80.7	84.7	94.7	69.3
200	95.0	76.5	85.5	63.5	74.0	77.0
240	84.6	66.3	87.9	55.4	63.3	78.3

Kiểm tra với tập Blance-Scale

Sự phân bố số mẫu trong ba lớp là 49, 288 và 288. Các kết quả ở Bảng 4.5 thể hiện rằng khả năng phân lớp đúng của Original EFART cao hơn đáng kể so với các mô hình tốt nhất trong mọi kiểm tra con (5%, 24.5%, 30.33%, 22.25%, và 3.2%), ngoại trừ kiểm tra con cuối cùng (Original EFART thấp hơn 5.9% so với mô hình tốt nhất - Original Fuzzy ART). Tuy nhiên, Complement EFART cao hơn trong mọi kiểm tra con. Đặc biệt, trong kiểm tra con cuối với số mẫu cao nhất cải thiện 4% so với mô hình tốt nhất thứ hai.

Kiểm tra với tập R15

Sự phân bố số mẫu trong 15 lớp là đồng đều mỗi lớp có 40 mẫu. Các số liệu từ Bảng 4.6 thể hiện rằng khả năng phân lớp đúng của Complement EFART là bằng với mô hình tốt nhất (Euclidean ART) trong bốn kiểm tra con cuối cùng và hơi thấp hơn trong hai kiểm tra con đầu.

Bảng 4.5: Kết quả phân lớp đúng của tập Blance-Scale

Số mẫu	OriEFART	ComEFART	OriFART	ComFART	EucART	K-mean
100	41	37	36	30	10	24
200	69	44.5	46	42	7.5	25
300	79.33	49	46.67	43.33	5	27.67
400	80	57.75	49.75	46.25	17	31.25
500	67	63.8	57.2	51.2	32.2	28.2
625	53.6	63.68	59.52	55.52	45.76	33.6

Bảng 4.6: Kết quả phân lớp đúng của tập R15

Số mẫu	OriEFART	ComEFART	OriFART	ComFART	EucART	K-mean
100	96.0	98.0	95.0	98.0	100	100
200	95.5	95.5	93.5	95.5	96.0	73.0
300	95.3	95.7	88.3	95.7	95.7	53.7
400	96.0	96.8	86.8	96.8	96.8	64.0
500	96.8	97.4	89.4	97.4	97.4	71.2
600	97.3	97.8	91.2	97.8	97.8	76.0

Kiểm tra với tập Glass

Sự phân bố số mẫu trong bảy lớp lần lượt là 70, 76, 17, 0, 13, 9, và 29. Đặc biệt số mẫu trong lớp 4 là 0. Bảng 4.7 thể hiện rằng khả năng phân lớp đúng của Original EFART cao hơn đáng kể so với mô hình tốt nhất trong ba kiểm tra con cuối cùng (15.3%, 13.5%, và 12.6) và thấp hơn khá nhiều trong hai kiểm tra con đầu.

Bảng 4.7: Kết quả phân lớp đúng của tập Glass

Số mẫu	OriEFART	ComEFART	OriFART	ComFART	EucART	K-mean
50	16.0	2.0	12.0	12.0	8.0	82.0
100	49.0	26.0	42.0	25.0	25.0	53.0
150	56.0	35.3	36.0	40.7	30.7	35.3
200	53.0	43.5	39.5	33.5	36.5	36.5
214	55.6	46.3	43.0	37.4	36.9	40.7

Kiểm tra với tập Wine

Sự phân bố số mẫu trong ba lớp lần lượt là 59, 71, và 48. Các kết quả từ Bảng 4.8 thể hiện rằng khả năng phân lớp đúng của Complement EFART xấp xỉ bằng mô hình tốt nhất – Kmean.

Bảng 4.8: Kết quả phân lớp đúng của tập Wine

Số mẫu	OriEFART	ComEFART	OriFART	ComFART	EucART	K-mean
30	100	100	100	73.3	100	100
60	98.3	98.3	98.3	68.3	98.3	100
90	83.3	88.9	85.6	64.4	66.7	90.0
120	76.7	84.2	82.5	60.8	50.0	86.7
150	77.3	85.3	82.7	64.7	40.7	85.3
178	77.5	87.6	83.7	69.7	34.3	87.6

Kiểm tra với tập Jain

Sự phân bố số mẫu trong hai lớp là 276 và 97. Các số liệu từ Bảng 4.9 thể hiện rằng khả năng phân lớp đúng của Complement EFART xấp xỉ bằng mô hình tốt nhất – Complement Fuzzy ART.

Bảng 4.9: Kết quả phân lớp đúng của tập Jain

Số mẫu	OriEFART	ComEFART	OriFART	ComFART	EucART	K-mean
100	99.0	100	99.0	100	100	100
200	99.5	100	99.5	100	57.0	100
300	96.3	97.7	69.7	100	43.0	100
373	94.6	94.4	69.2	99.7	47.5	97.9

Kiểm tra với tập Aggregation

Sự phân bố số mẫu trong bảy lớp lần lượt là 45, 170, 102, 273, 34, 130 và 34. Bảng 4.10 thể hiện rằng khả năng phân lớp đúng của Complement EFART hơi thấp hơn mô hình tốt nhất – Euclidean ART trong ba kiểm tra con đầu nhưng lại cao hơn 4.8 % trong kiểm tra con cuối cùng

Bảng 4.10: Kết quả phân lớp đúng của tập Aggregation

Số mẫu	OriEFART	ComEFART	OriFART	ComFART	EucART	K-mean
200	98.0	96.5	83.5	83.5	100	81.5
400	88.8	91.8	68.3	82.0	98.0	66.3
600	83.7	93.0	59.5	84.8	95.5	65.2
788	69.2	78.0	51.3	68.3	73.2	52.9

Kết quả từ các kiểm tra con của 9 thử nghiệm được tổng hợp trong Bảng 4.1. Các số liệu cho thấy sự cải thiện khả năng phân lớp của EFART so với mô hình tốt nhất thứ hai.

Dữ liệu của Bảng 4.11 thể hiện rằng EFART thích hợp cho các tập dữ liệu nhỏ, phức tạp. EFART phân lớp tốt nhất với tập dữ liệu có các đặc trưng sau: sự phân bố số mẫu tại các lớp là không đều với độ lệch cao, số lượng lớp là nhỏ/trung bình, số lượng thuộc tính là nhỏ/trung bình, và số lượng mẫu là nhỏ/trung bình.

Bảng 4.11: Sự cải thiện khả năng phân lớp của EFART với luật học thứ nhất so với mô hình tốt nhất thứ hai

Kiểu tập dữ liệu	Sự phân bố số mẫu trong các lớp	Số mẫu	Số lớp	Số thuộc tính	Mức độ cải thiện (%)
1	Không đều với độ lệch cao	200-400	2 & 3	2	21-32.6
2	Không đều với độ lệch trung bình	150	3	9	15.3
3	Không đều với độ lệch trung bình	200-214	7	9	12.6-13.5
4	Không đều với độ lệch cao	500	3	2	9.8
5	Không đều với độ lệch cao	788	7	7	4.8
6	Không đều với độ lệch thấp	120-200	2 & 3	2 & 4	2.5-5.3
7	Đều với mọi phân lớp	90-201	2	3 & 4	1.5-5
8	Không đều với độ lệch trung bình	250	3	2	0.8

4.5.2 Thử nghiệm 2: Dùng luật học thứ hai

7 tập dữ liệu chuẩn được chọn từ cơ sở dữ liệu UCI bao gồm MONKS, BALANCE-SCALE, D31, R35, WDBC (Wisconsin Diagnostic Breast Cancer), WINE-RED (Wine Quality of Red wine), and WINE-WHITE (Wine Quality of White wine). Các tập dữ liệu này khác nhau số thuộc tính, số lớp, số mẫu huấn luyện, và sự phân bố các mẫu ở các lớp. Bảng 4.12 thể hiện các thông tin trên của các tập dữ liệu được chọn.

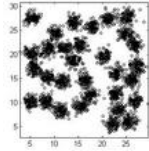
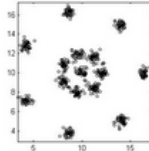
Kiểm tra với tập WDBC

Sự phân bố số mẫu trong hai lớp là không đều với mức độ chênh lệch trung bình. Dữ liệu từ Bảng 4.13 cho thấy khả năng phân lớp đúng của Complement EFART là cao hơn đáng kể so với các mô hình khác trong mọi kiểm tra con.

Kiểm tra với tập D31

Sự phân bố số mẫu trong 31 lớp là đồng đều. Bảng 4.14 cho thấy Complement EFART là mô hình tốt nhất trong mọi kiểm tra con.

Bảng 4.12: Đặc trưng của các tập dữ liệu trong thử nghiệm 2

Thứ tự	Tên tập dữ liệu	Số lớp	Số thuộc tính	Số mẫu	Nội dung dữ liệu
1	WDBC	2	30	569	Chuẩn đoán bệnh ung thư
2	WINE-RED	6	11	1599	Chất lượng rượu vang đỏ
3	WINE-WHITE	6	11	4898	Chất lượng rượu vang trắng
4	BALANCE-SCALE	3	4	625	Thử nghiệm tâm lý học
5	MONKS	2	6	459	Các bài toán của Monk
6	D31	31	2	3100	 Các điểm dữ liệu tạo hình
7	R15	15	2	600	 Các điểm dữ liệu tạo hình

Kiểm tra với tập WINE-WHITE

Sự phân bố số mẫu trong sáu lớp là không đều với mức độ chênh lệch cao. Số liệu của Bảng 4.15 cho thấy khả năng phân lớp đúng của Original EFART là cao hơn đáng kể so với các mô hình khác trong mọi kiểm tra con

Bảng 4.13: Kết quả phân lớp đúng của tập WDBC

Số mẫu	OriEFART	ComEFART	OriFART	ComFART	EucART	K-mean
569	55.89	90.51	35.85	74.17	46.92	16.17
500	51	89.2	36.4	73.6	41.8	18.4
400	40	86.5	39	70.75	33.75	23
300	26	82	32.33	67.33	21	30.67
200	0	92.5	21.5	76	0.5	44.5
100	0	89	8	54	0	45

Bảng 4.14: Kết quả phân lớp đúng của tập D31

Số mẫu	OriEFART	ComEFART	OriFART	ComFART	EucART	K-mean
3100	91.87	94.45	84.74	92.94	92.48	65
2500	91.44	94.16	84.44	91.96	91.64	64.68
2000	90.7	93.95	86.1	91.55	91.2	61.6
1500	90	93.87	85.67	91.6	90.8	59.93
1000	89.5	93.2	86.4	90.9	90.5	66.2
500	84.8	90.6	89.6	86.8	88.6	57.4

Bảng 4.15: Kết quả phân lớp đúng của tập WINE-WHITE

Số mẫu	OriEFART	ComEFART	OriFART	ComFART	EucART	K-mean
4898	34.73	43.32	21.95	17.78	17.07	32.24
4000	37.75	33.18	23.48	15.425	18.45	32.1
3000	41.57	12.67	28.47	18.13	16.93	30.47
2000	43.6	4.3	26.3	21.3	19.25	29.8
1000	50.8	4.1	11	23.8	23.7	20.7

Kiểm tra với tập BALANCE-SCALE

Sự phân bố số mẫu trong ba lớp là không đều với mức độ chênh lệch cao. Bảng 4.16 cho thấy Original EFART phân lớp tốt hơn đáng kể so với các mô hình khác trong mọi kiểm tra con ngoại trừ kiểm tra cuối cùng với số mẫu ít nhất (100 trong số 625 mẫu).

Kiểm tra với tập R15

Sự phân bố số mẫu trong 15 lớp là đồng đều. Dữ liệu của Bảng 4.17 cho thấy kết quả phân lớp đúng của Original EFART cao hơn so với các mô hình khác trong mọi kiểm tra con ngoại trừ kiểm tra cuối cùng với số mẫu ít nhất (100 trong số 600 mẫu) .

Kiểm tra với tập MONK

Sự phân bố số mẫu trong hai lớp là đồng đều. Bảng 4.18 cho thấy Original EFART phân lớp tốt hơn so với các mô hình khác trong mọi kiểm tra con ngoại trừ kiểm tra cuối cùng với số mẫu ít nhất (100 trong số 459 mẫu) .

Kiểm tra với tập WINE-RED

Sự phân bố số mẫu trong sáu lớp là không đều với mức độ chênh lệch cao. Số liệu của Bảng 4.19 cho thấy khả năng phân lớp đúng của Original EFART là tốt hơn so với các mô hình khác trong mọi kiểm tra con nhưng hơi thấp hơn Complement Fuzzy ART trong kiểm tra con đầu tiên (0.13%).

Kết quả từ các thử nghiệm con của 7 thử nghiệm được tổng hợp trong Bảng 4.20. Các dữ liệu thể hiện sự cải thiện khả năng phân lớp của EFART so với mô hình tốt nhất thứ hai.

Bảng 4.20 cho thấy khả năng của phân lớp của Fuzzy ART với luật học thứ hai cải thiện đáng kể đối với các tập dữ liệu nhỏ có độ phức tạp cao (nhiều thuộc tính, phân phối số mẫu không đồng đều với độ lệch cao). Đặc biệt, kết quả phân lớp là cao khi tập dữ liệu có chứa nhiều mẫu

Bảng 4.16: Kết quả phân lớp đúng của tập BALANCE-SCALE

Số mẫu	OriEFART	ComEFART	OriFART	ComFART	EucART	K-mean
625	80.16	67.52	59.52	55.52	45.76	33.6
500	75.2	64.8	57.2	51.2	32.2	28.2
400	70.5	56.25	49.75	46.25	17	31.25
300	67	46.67	46.67	43.33	5	27.67
200	50.5	44	46	42	7.5	25
100	19	38	36	30	10	24

Bảng 4.17: Kết quả phân lớp đúng của tập R15

Số mẫu	OriEFART	ComEFART	OriFART	ComFART	EucART	K-mean
600	98.17	98.17	91.2	97.8	97.8	76
500	97.8	97.8	89.4	97.4	97.4	71.2
400	97.25	97.25	86.8	96.8	96.8	64
300	96.33	96.33	88.3	95.7	95.7	53.7
200	96.5	96.5	93.5	95.5	96	73
100	99	99	95	98	100	100

Bảng 4.18: Kết quả phân lớp đúng của tập MONK

Số mẫu	OriEFART	ComEFART	OriFART	ComFART	EucART	K-mean
459	67.97	41.18	42.92	41.18	65.36	45.75
400	63.25	41.5	44.25	42	60.75	45.25
300	57.67	42.67	44	37.67	47.67	48.67
200	64.5	46	48	41.5	23	59.5
100	78	65	59	52	2	89

Bảng 4.19: Kết quả phân lớp đúng của tập WINE-RED

Số mẫu	OriEFART	ComEFART	OriFART	ComFART	EucART	K-mean
1599	25.39	17.32	18.26	25.52	14.26	16.77
1200	33.83	23.08	20.25	18.67	17.75	17
900	41	27.78	21.11	10.78	22	19.22
600	32	26.83	21.5	9.833	28.67	17.83
300	12.67	21.33	25.67	16.33	51.67	24.67

Bảng 4.20: Sự cải thiện khả năng phân lớp của EFART so với mô hình tốt nhất thứ hai trong thử nghiệm 2

Kiểu tập dữ liệu	Sự phân bố số mẫu trong các lớp	Số mẫu	Số lớp	Số thuộc tính	Mức độ cải thiện (%)
1	Không đều với độ lệch cao	600-1200	3,4	11	15.2-30.2
2	Không đều với độ lệch cao	1000-3000	3,4	11	11.1-30.1
3	Không đều với độ lệch cao	300-625	2,3	4	18-20.8
4	Không đều với độ lệch cao	400-569	2	30	15.6-16.3
5	Không đều với độ lệch cao	4000-4898	5,6	11	2.5-5.7
6	Không đều với độ lệch trung bình	400-459	2	7	2.5-2.6
7	Đều với mọi phân lớp	500-1500	5,10,15	2	2.3-3.8
8	Đều với mọi phân lớp	2000-3100	20,25,31	2	1.5-2.4

Cả hai thử nghiệm cho thấy Fuzzy ART với các luật học đề xuất cải thiện đáng kể khả năng phân lớp cho các tập dữ liệu nhỏ, phức tạp. Tuy nhiên, K-mean và Euclidean ART phân lớp tốt hơn đối với các tập dữ liệu có mật độ cao các mẫu co lại thành lớp và đường biên giữa các lớp rõ ràng.

4.6 Kết luận chương

Mô hình Fuzzy ART là một ANN thực hiện lưu dữ liệu theo các cụm với một số ưu điểm nổi trội. Các ưu điểm này gồm: học dữ liệu huấn luyện cho đến khi thỏa mãn điều kiện nhất định; có thể sinh ra các phân cụm mới nhưng không ảnh hưởng đến các phân cụm hiện có; dễ dàng lựa chọn các tham số cho mô hình dựa vào các nghiên cứu về tính chất toán học của tập mờ. Tuy nhiên, các luật học của các mô hình Fuzzy ART học chưa hiệu quả các mẫu huấn luyện. Điều này dẫn đến khả năng nhớ lại các cụm của các Fuzzy ART này bị hạn chế. Do đó, việc cải tiến hay phát triển các luật học hiệu quả cho Fuzzy ART là hết sức cần thiết để tăng khả năng phân cụm của các Fuzzy ART.

Hai luật học hiệu quả được đề xuất để học các mẫu huấn luyện tốt hơn. Hai luật này học mọi mẫu huấn luyện và giảm ảnh hưởng của các mẫu huấn luyện dị bộ. Một thuật toán xác định giá trị thích hợp của tham số học được trình bày. Các thử nghiệm với 14 tập dữ liệu chuẩn được làm để chứng minh tính hiệu quả của hai luật học đề xuất. Kết quả thực nghiệm cho thấy Fuzzy ART gắn với hai luật học cải thiện đáng kể chất lượng phân cụm đối với các tập dữ liệu nhỏ, phức tạp. Kết quả nghiên cứu này được công bố tại tạp chí Journal of Information Science and Engineering với chỉ số SCIE (Công trình khoa học số 10), Tạp chí Khoa học & Công nghệ trong nước (Công trình khoa học số 8), kỷ yếu hội thảo có phản biện của Hội nghị quốc tế của IEEE về Công nghệ tính toán và truyền thông – IEEE RIVF 2013 (Công trình khoa học số 4), và phiên Quốc tế của Hội nghị thường niên lần thứ 17 của Đài Loan về công nghệ và ứng dụng của Trí tuệ nhân tạo – TAAI 2012 (Công trình khoa học số 1).

CHƯƠNG 5. LUẬT HỌC CẢI TIẾN CHO BỘ NHỚ LIÊN KẾT MỜ

Trong phần 2.5 của chương 2, tác giả đã trình bày các hiểu biết quan trọng về FAM. Trong chương này, tác giả sẽ trình bày các nghiên cứu có liên quan đến FAM để làm cơ sở lý luận đề xuất cải tiến quá trình học cho FAM. Tiếp theo, mô hình FAM với luật học cải tiến được trình bày. Trong Phần 5.5 các kết quả thực nghiệm được mô tả và phân tích chi tiết.

5.1 Giới thiệu chung

FAM là AM cho phép học và nhớ lại với các mẫu được biểu diễn ở dạng tập mờ. FAM có ưu điểm quan trọng là nhớ lại các mẫu đã lưu từ các mẫu vào nhiều hay không đầy đủ. Do các mẫu được thể hiện ở dạng tập mờ nên FAM được áp dụng trong việc suy diễn, xử lý với các thông tin không chính xác. Vì vậy, FAM được áp dụng cho rất nhiều ứng dụng trong lĩnh vực như xử lý ảnh, ước lượng và dự báo.

5.2 Các nghiên cứu liên quan

Các nghiên cứu về FAM có thể chia thành hai nhóm chính gồm phát triển các mô hình mới và áp dụng mô hình lý thuyết vào các ứng dụng thực. Ngoài ra, có một số lượng nhỏ các nghiên cứu khác về lý thuyết như các điều kiện để nhớ lại đúng các mẫu đã lưu và chuyển đổi thành các FAM mới từ một FAM ban đầu

5.2.1 Các mô hình lý thuyết

Trong hướng nghiên cứu này, các nhà khoa học sử dụng các thao tác của toán học, logic mờ và toán học hình thái trong quá trình học và nhớ lại để tạo ra các FAM chuẩn và một số biến thể của FAM.

Các FAM chuẩn được nghiên cứu rộng rãi do ba ưu điểm của chúng về tính chịu nhiễu, khả năng lưu trữ và tính hội tụ. Kosko [43,44] đưa ra mô hình FAM đầu tiên lưu trữ luật mờ “Nếu X là X_k thì Y là Y_k ” bằng cách sử dụng luật Hebb mờ. Kosko đưa ra hai cách tổng hợp ma trận trọng số là max-min và tích cực đại. Tuy

nhiên, khả năng lưu trữ lại thấp do lưu mỗi luật trong một ma trận nên hệ thống mờ của Kosko gồm nhiều ma trận và kết quả ra của hệ thống được tổng hợp từ kết quả ra của các ma trận thể hiện một luật mờ đó. Junbo và đồng nghiệp [40] đưa ra một luật học cho FAM với sự tổng hợp max-min trong quá trình nhớ lại. Với một số điều kiện, luật đưa ra có thể mã hóa hiệu quả nhiều cặp mẫu trong một FAM đơn và thu được các liên kết hoàn hảo. Chung và Lee [14,15] cũng đưa ra một mô hình FAM và một định lý về nhớ lại hoàn hảo các mẫu đã lưu. Tiếp theo, hai thuật toán mã hóa hiệu quả (mã hóa trực giao, mã hóa trọng số) được công bố. Sau đó, hai tác giả này còn đưa ra kết quả nghiên cứu về việc thêm vào hay bớt đi một mẫu trong FAM [15]. Xiao và đồng nghiệp [71] đưa ra một thuật toán học mã hóa max-min cho FAM với việc lấy cực đại của phép nhân giữa mẫu vào và ma trận trọng số để nhớ lại. Phương thức mới này có thể lưu trữ các mẫu của bộ nhớ liên kết trong chế độ tự liên kết. Mã hóa max-min được dùng để tính toán trọng số kết nối của FAM. S.T.Wang và Lu [65] đã thiết kế một tập các FAM dựa vào các thao tác của logic mờ và toán học hình thái với hai thao tác chính là giãn nở và co rút. Các FAM với thao tác giãn nở làm việc tốt với các nhiều dạng giãn nở. Còn các FAM với thao tác co rút làm việc tốt với các nhiều dạng co rút. Tập FAM này ở chế độ tự liên kết có khả năng lưu trữ không giới hạn và hội tụ trong một lần lặp. Nhóm các FAM này có thể làm hiệu quả với các nhiều pha trộn. Sussner và Valle [58,59] giới thiệu một họ các FAM gợi ý. Mỗi FAM gợi ý gồm một mạng các nơ-ron Pedrycz logic kết nối trong. Các trọng số liên kết được xác định bởi cực tiểu của các gợi ý của mỗi cặp mẫu. Họ mô hình này cũng có ưu điểm về khả năng lưu trữ không giới hạn, hội tụ sau một lần lặp và chịu nhiễu tốt với các mẫu co rút. Hơn nữa hai tác giả còn trình bày về một số ưu điểm của họ mô hình này và thảo luận mối quan hệ giữa FAM gợi ý và Bộ nhớ liên kết hình thái (MAM – *Morphological Associative Memory*).

5.2.2 Các biến thể của FAM

Một số dạng biến thể của FAM cũng được công bố với một số ít các nghiên cứu. Li và đồng nghiệp [47] đưa ra một FAM trực giác. Dựa vào thao tác gợi ý mờ của Godel, một luật học cho nhiều cặp mẫu trực giác cho FAM được đưa ra. Với một số điều kiện nhất định, mô hình này mã hóa hiệu quả các mẫu trong một FAM

đơn và thu được các liên kết hoàn hảo. Mô hình này thêm vào các giá trị bù cho mỗi phân tử của mẫu huấn luyện. Valle [62] trình bày một giải pháp cho FAM khi nhớ lại từ các mẫu vào có nhiều pha trộn bằng cách giới thiệu một lớp FAM gợi ý dựa vào sự hoán đổi. Các mô hình này cũng có ưu điểm về sự hội tụ và khả năng nhớ trong chế độ tự liên kết. Esmi và Susner [21] dùng ưu điểm của thao tác co rút của toán học hình thái để đưa ra một FAM không phân tán dựa vào thước đo tập con láng riêng (subsethood) của Kosko. Ngoài ra, Zeng và đồng nghiệp [75] thiết lập một khái niệm mới về điểm mạnh của FAM khi làm việc với các mẫu xáo trộn. Các tác giả cũng chỉ ra FAM với suy diễn tích cực đại là tốt khi dùng luật học của Hebb nhưng sẽ không hiệu quả với luật học khác.

5.2.3 Một số mô hình FAM

Các FAM có đặc điểm là cả quá trình học và nhớ lại chỉ thực hiện trong một lần lặp bằng việc dùng các phép thao tác của toán học, logic mờ và toán học hình thái.

Giả sử các FAM lưu p cặp mẫu $(\mathbf{A}^k, \mathbf{B}^k)$ với \mathbf{A}^k có n phần tử và \mathbf{B}^k có m phần tử.

Mô hình của Kosko và Kong [43]

Mô hình FAM này sử dụng các phép toán lấy cực tiểu để học các liên kết giữa mỗi cặp mẫu và dùng phép toán giãn nở để tổng quát hóa các liên kết và tính toán mẫu ra.

Các ma trận trọng số kết nối \mathbf{W}^k của cặp mẫu $(\mathbf{A}^k, \mathbf{B}^k)$ được tính bằng công thức:

$$W_{ij}^k = \min(A_i^k, B_j^k) \quad (5.1)$$

Ma trận trọng số kết nối \mathbf{W} lưu tất cả các cặp mẫu được tính bằng công thức:

$$W_{ij} = \bigvee_{k=1}^p W_{ij}^k \quad (5.2)$$

Với mẫu vào \mathbf{X} , mẫu ra \mathbf{Y} được tính bằng hàm đầu ra:

$$Y_j = \max \left\{ \bigwedge_{i=1}^m (X_i, W_{ij}) \right\} \quad (5.3)$$

Mô hình của Junbo, Fan, và Yan [40]

Mô hình FAM này sử dụng thao tác gợi ý mờ được dùng để trình bày liên kết giữa các mẫu. Ma trận trọng số lưu các liên kết được tính toán bởi thao tác co rút và sử dụng lại hàm đầu ra của Kosko.

Các ma trận trọng số kết nối \mathbf{W}^k của cặp mẫu $(\mathbf{A}^k, \mathbf{B}^k)$ được tính bằng công thức:

$$W_{ij}^k = I_M(A_i^k, B_j^k) \quad (5.4)$$

Ma trận trọng số kết nối \mathbf{W} lưu tất cả các cặp mẫu được tính bằng công thức:

$$W_{ij} = \bigwedge_{k=1}^p W_{ij}^k \quad (5.5)$$

Với mẫu vào \mathbf{X} , mẫu ra \mathbf{Y} được tính bằng hàm đầu ra:

$$Y_j = \max \left\{ \bigwedge_{i=1}^m (X_i, W_{ij}) \right\} \quad (5.6)$$

Mô hình của Chung và Lee [14]

Hai mô hình FAM này thiết kế giống như mô hình của Junbo nhưng hàm đầu ra sử dụng một thao tác *t-norm* để tính toán.

Các ma trận trọng số kết nối \mathbf{W}^k của cặp mẫu $(\mathbf{A}^k, \mathbf{B}^k)$ được tính bằng công thức:

$$W_{ij}^k = I_M(A_i^k, B_j^k) \quad (5.7)$$

Ma trận trọng số kết nối \mathbf{W} lưu tất cả các cặp mẫu được tính bằng công thức:

$$W_{ij} = \bigwedge_{k=1}^p W_{ij}^k \quad (5.8)$$

Với mẫu vào \mathbf{X} , mẫu ra \mathbf{Y} được tính bằng hàm đầu ra:

$$Y_j = \bigvee_{i=1}^m \varphi(X_i, W_{ij}) \quad (5.9)$$

với φ là một *t-norm*

Mô hình của Xiao, Yang, và Yu [71]

Các tác giả thiết kế một mô hình mới có sử dụng tỷ lệ giữa các mẫu vào và các mẫu ra để lưu các liên kết giữa các mẫu và dùng phép toán co rút để tổng quát hóa.

Các ma trận trọng số kết nối \mathbf{W}^k của cặp mẫu $(\mathbf{A}^k, \mathbf{B}^k)$ được tính bằng công thức:

$$W_{ij}^k = \varphi(A_i^k, B_j^k) \quad (5.10)$$

với φ được tính như sau:

$$\varphi(x, y) = \begin{cases} 1, x = y \\ \frac{\min(A_i^k, B_j^k)}{\max(A_i^k, B_j^k)}, x \neq y \end{cases} \quad (5.11)$$

Ma trận trọng số kết nối \mathbf{W} lưu tất cả các cặp mẫu được tính bằng công thức:

$$W_{ij} = \bigwedge_{k=1}^p W_{ij}^k \quad (5.12)$$

Với mẫu vào \mathbf{X} , mẫu ra \mathbf{Y} được tính bằng hàm đầu ra:

$$Y_j = \bigvee_{i=1}^m X_i \cdot W_{ij} \quad (5.13)$$

Tập các mô hình của S.T. Wang và Lu [65]

Tập các FAM này lưu trữ liên kết của các cặp giống như Fulai nhưng việc tổng quát hóa các liên kết và tính toán mẫu ra lại có thể làm bằng một trong hai phép toán là co rút hay giãn nở.

Các ma trận trọng số kết nối \mathbf{W}^k của cặp mẫu $(\mathbf{A}^k, \mathbf{B}^k)$ được tính bằng công thức:

$$W_{ij}^k = B_i^k / A_j^k \quad (5.14)$$

Ma trận trọng số kết nối \mathbf{W} lưu tất cả các cặp mẫu được tính bằng công thức:

$$W_{ij} = \bigwedge_{k=1}^p W_{ij}^k \quad (5.15)$$

hoặc

$$W_{ij} = \bigvee_{k=1}^p W_{ij}^k \quad (5.16)$$

Với mẫu vào \mathbf{X} , mẫu ra \mathbf{Y} được tính bằng hàm đầu ra:

$$Y_j = \bigvee_{i=1}^m (X_i + W_{ij}) \quad (5.17)$$

hoặc

$$Y_j = \bigwedge_{i=1}^m (X_i + W_{ij}) \quad (5.18)$$

Họ các FAM gợi ý của Sussner và Valle [58]

Hai tác giả dùng phép toán gợi ý mờ để học các liên kết giữa các mẫu. Sau đó các liên kết được tổng quát hóa bằng một trong hai thao tác là co rút hay giãn nở. Để tính toán mẫu ra, một thao tác s -norm đã được sử dụng và một ngưỡng được thêm vào để cải thiện mẫu ra của FAM.

Các ma trận trọng số kết nối \mathbf{W}^k của cặp mẫu $(\mathbf{A}^k, \mathbf{B}^k)$ được tính bằng công thức:

$$W_{ij}^k = \varphi(A_i^k, B_j^k) \quad (5.19)$$

với φ là thao tác I_M, I_P, I_L

Ma trận trọng số kết nối \mathbf{W} lưu tất cả các cặp mẫu được tính bằng công thức:

$$W_{ij} = \bigwedge_{k=1}^p W_{ij}^k \quad (5.20)$$

hoặc

$$W_{ij} = \bigvee_{k=1}^p W_{ij}^k \quad (5.21)$$

Với mẫu vào \mathbf{X} , mẫu ra \mathbf{Y} được tính bằng hàm đầu ra:

$$Y_j = \psi(X_i, W_{ij}) \vee \theta_j \quad (5.22)$$

với ψ là thao tác D_M, D_P, D_L và θ được tính bằng công thức:

$$\theta_j = \bigwedge_{k=1}^p B_j^k \quad (5.23)$$

Các mô hình FAM ở trên chỉ thể hiện tốt nội dung hoặc liên kết giữa các cặp mẫu nên không thể hiện tốt cả nội dung và liên kết của các mẫu.

5.3 Lý do đề xuất luật học cải tiến cho FAM

Các mô hình FAM có thể chia thành hai nhóm chính. Nhóm thứ nhất gồm các mô hình FAM chỉ lưu trữ tốt thành phần của các mẫu được học. Liên kết của mỗi cặp mẫu được thể hiện bằng giá trị cực tiểu của hai phần tử tương ứng trong mỗi cặp mẫu [71] hay bằng giá trị một phần tử của một mẫu trong cặp mẫu khi thỏa mãn điều kiện sau: giá trị phần tử của mẫu vào có giá trị nhỏ hơn giá trị phần tử của mẫu ra [14,40,58]. Tuy nhiên, các FAM này không lưu trữ hiệu quả sự liên kết giữa mẫu vào và mẫu ra. Trong nhóm hai, các FAM chỉ lưu trữ tốt liên kết giữa các mẫu. Bộ nhớ cơ bản của mỗi cặp mẫu lưu tỷ lệ giữa mẫu vào và mẫu ra [58,65,71]. Tuy nhiên, thành phần của các mẫu lại không được lưu trữ hiệu quả trong các FAM này.

Từ các phân tích trên, tác giả thấy rằng một số thông tin có ích sẽ bị mất mát trong quá trình học do các mô hình FAM chỉ lưu trữ tốt nội dung các mẫu hoặc chỉ lưu trữ tốt liên kết giữa các mẫu. Thông tin bị mất mát trong quá trình học có thể dẫn tới khả năng nhớ lại bị hạn chế. Do đó, luật học cải tiến cho FAM được đề xuất để lưu trữ tốt cả nội dung và sự liên kết của các mẫu. Từ việc cải tiến quá trình học, khả năng chịu nhiễu của FAM có thể sẽ tăng lên.

5.4 Luật học cải tiến

5.4.1 Ý tưởng

Mục đích: Cải thiện khả năng lưu trữ của FAM để FAM lưu trữ hiệu quả cả nội dung và sự liên kết của các mẫu. Từ việc lưu trữ tốt các mẫu, FAM sẽ cải thiện khả năng nhớ lại từ các mẫu vào nhiều.

Ý tưởng: Để lưu trữ nội dung của các mẫu trong FAM, nội dung của mẫu ra được dùng. Sự chênh lệch giữa nội dung của mẫu vào và mẫu ra được dùng để lưu liên kết giữa hai mẫu trong cặp. Sự tổng hợp của mẫu ra và sự chênh lệch giữa mẫu ra và mẫu vào được lưu trữ trong FAM.

5.4.2 Mô hình FAM với luật học cải tiến

Mô hình FAM với luật học cải tiến (ACAM - *Association-Content Associative Memory*) được xây dựng dựa trên các bộ nhớ liên kết mờ gợi ý (IFAM) [58]. Cụ thể

hơn, các phép toán mờ của Lukasiewicz được dùng gồm phép nối (C_L), phép chia tách (D_L) và phép gợi ý (I_L).

Giả sử FAM học p cặp mẫu gồm $(\mathbf{A}^1, \mathbf{B}^1), \dots, (\mathbf{A}^p, \mathbf{B}^p)$. Mỗi cặp mẫu $(\mathbf{A}^k, \mathbf{B}^k)$ được lưu trong bộ nhớ cơ bản \mathbf{W}^k . Sự tổng quát hóa p cặp mẫu được lưu trữ trong ma trận trọng số chung \mathbf{W} . Cặp mẫu thứ k được biểu diễn bởi hai véc tơ: $\mathbf{A}^k = (A_1^k, \dots, A_n^k)$ và $\mathbf{B}^k = (B_1^k, \dots, B_m^k)$.

Thiết kế của mô hình FAM với luật học cải tiến được trình bày như sau:

Quá trình học các cặp mẫu

Việc học và lưu trữ p cặp mẫu trong FAM được thực hiện qua hai bước sau:

Bước 1: Học và lưu trữ cặp mẫu $(\mathbf{A}^k, \mathbf{B}^k)$ trong ma trận trọng số \mathbf{W}^k theo cách sau:

$$W_{ij}^k = \eta(B_j^k - A_i^k) + (1 - \eta)B_j^k = B_j^k - \eta A_i^k \quad (5.24)$$

với η là nhân tố điều khiển tỷ lệ giữa nội dung và sự liên kết được lưu. Do dùng nhân tố η nên khi mẫu vào bị nhiễu thì nhiễu sẽ ít ảnh hưởng hơn đến mẫu được nhớ lại.

Bước 2: Tổng quát hóa sự liên kết của các cặp mẫu và lưu trữ trong ma trận trọng số chung theo công thức sau:

$$W_{ij} = \bigwedge_{k=1}^p W_{ij}^k = \bigwedge_{k=1}^p (B_j^k - \eta A_i^k) \quad (5.25)$$

Quá trình nhớ lại

Việc nhớ lại mẫu ra \mathbf{Y} từ mẫu vào \mathbf{X} và ma trận trọng số chung \mathbf{W} được thực hiện như sau: Mẫu ra được nhớ lại thông qua sự tổng hợp của ma trận trọng số chung và mẫu vào như trong công thức dưới đây:

$$Y_j = \bigvee_{i=1}^m (\eta X_i + W_{ij}) \quad (5.26)$$

Điều chỉnh ACAM cho chế độ tự liên kết

Để đảm bảo khả năng lưu trữ không giới hạn các cặp mẫu ở chế độ tự liên kết, tác giả tính W_{ii} giống như trong MAM [52].

$$W_{ij} = \begin{cases} \bigwedge_{k=1}^p (B_j^k - A_i^k) & \text{if } i = j \\ \bigwedge_{k=1}^p (B_j^k - \eta A_i^k) & \text{if } i \neq j \end{cases} \quad (5.27)$$

Khi đó, công thức dùng cho việc nhớ lại được điều chỉnh như sau:

$$Y_j = \bigvee_{i \neq j} (\eta X_i + W_{ij}) \vee (X_i + W_{ij}) \quad (5.28)$$

5.4.3 Định lý và hệ quả về khả năng nhớ lại hoàn hảo của FAM cải tiến

Trong phần này, điều kiện nhớ lại hoàn hảo các mẫu đã lưu của ACAM ở chế độ liên kết khác loại được nêu và chứng minh trong định lý thứ nhất. Tiếp theo, khả năng nhớ lại đúng mọi cặp mẫu đã lưu của ACAM ở chế độ tự liên kết được khẳng định và chứng minh trong Hệ quả 5.1.

Định lý 5.1: (Khả năng nhớ lại các cặp mẫu trong chế độ liên kết khác loại)

W trong công thức (5.27) nhớ lại hoàn hảo mọi cặp mẫu $(\mathbf{A}^k, \mathbf{B}^k)$ nếu và chỉ nếu với mỗi $k=1, \dots, p$, mỗi cột của ma trận $\mathbf{W}^k - \mathbf{W}$ có chứa một số 0.

Chứng minh: W nhớ lại hoàn hảo mọi cặp mẫu $(\mathbf{A}^k, \mathbf{B}^k)$ tương đương với

$$\bigvee_{i \neq j} (\eta A_i^k + W_{ij}) \vee (A_i^k + W_{ij}) = B_j^k$$

với mọi $j=1, \dots, m$

$$\Leftrightarrow B_j^k - \bigvee_{i \neq j} (\eta A_i^k + W_{ij}) \vee (A_i^k + W_{ij}) = 0$$

với mọi $k=1, \dots, p$ và với mọi $j=1, \dots, m$

$$\Leftrightarrow B_j^k + \bigwedge_{i \neq j} (-\eta A_i^k - W_{ij}) \wedge (-A_i^k - W_{ij}) = 0$$

với mọi $k=1, \dots, p$ và với mọi $j=1, \dots, m$

$$\Leftrightarrow \bigwedge_{i \neq j} (B_j^k - \eta A_i^k - W_{ij}) \wedge (B_j^k - A_i^k - W_{ij}) = 0$$

với mọi $k=1, \dots, p$ và với mọi $j=1, \dots, m$

$$\Leftrightarrow \bigwedge_{i=1}^n (W_{ij}^k - W_{ij}) = 0$$

với mọi $k=1, \dots, p$ và với mọi $j=1, \dots, m$

Cụm cuối cùng của công thức này là đúng khi và chỉ khi với mỗi $k=1, \dots, p$ và mỗi số nguyên $j=1, \dots, m$ thì một hàng trong cột thứ j của $[\mathbf{W}^k \ \mathbf{W}]$ chứa ít nhất một số 0.

Hệ quả 5.1: (Khả năng nhớ lại các cặp mẫu trong chế độ tự liên kết)

\mathbf{W} trong công thức (5.27) nhớ lại hoàn hảo mọi cặp mẫu $(\mathbf{A}^k, \mathbf{A}^k)$.

Chứng minh: bởi vì $W_{jj}^k = A_j^k - A_j^k = 0$ với mỗi $j=1, \dots, m$ và với mọi $k=1, \dots, p$ nên với mỗi $k=1, \dots, p$ thì mỗi cột của $\mathbf{W}^k \ \mathbf{W}$ chứa một số 0. Theo định lý 1 thì \mathbf{W} nhớ lại hoàn hảo mọi cặp mẫu $(\mathbf{A}^k, \mathbf{A}^k)$.

3.5 Kết quả thực nghiệm

Để so sánh tính hiệu quả của ACAM trong việc điều khiển nhiễu tốt hơn các bộ nhớ liên kết chuẩn khác, tác giả đã tiến hành các thực nghiệm. Năm mô hình được thực thi gồm các mô hình được đưa ra bởi Junbo [40], Kosko [44], Xiao [71], Ritter [52] (MAM), Susners và Valle (IFAMs) [58].

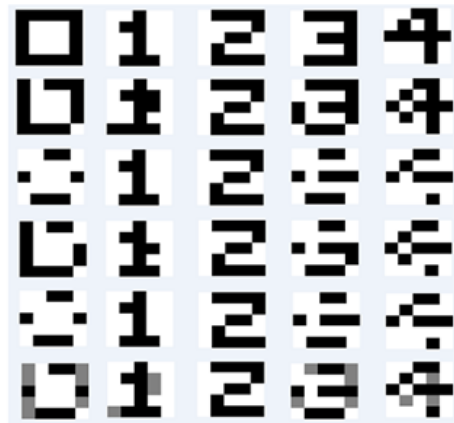
Để đánh giá kết quả, tác giả dùng cách tính lỗi chuẩn theo công thức sau:

$$E(\tilde{\mathbf{B}}, \mathbf{B}) = \frac{\|\tilde{\mathbf{B}} - \mathbf{B}\|}{\|\mathbf{B}\|} \quad (5.29)$$

với \mathbf{B} là mẫu ra mong đợi, $\tilde{\mathbf{B}}$ là mẫu ra từ quá trình nhớ lại và $\|\cdot\|$ là chuẩn L_2 của vector.

3.5.1 Thử nghiệm với tập dữ liệu về các số

Bộ dữ liệu này bao gồm 5 hình ảnh của số 0-4 với kích thước 5×5 . Mỗi ảnh được chuyển đổi thành một vector kích thước 1×25 bằng phương pháp quét theo hàng. Với dữ liệu này, kích thước của các ma trận trọng số \mathbf{W} là 25×25 được sử dụng để lưu trữ 5 mẫu có kích thước 1×25 . Tác giả thực hiện thí nghiệm với mẫu vào bị biến đổi với cả hai chế độ (tự liên kết và liên kết khác loại). Các ảnh bị biến đổi chứa cả nhiễu giãn nở và nhiễu co rút (nhiều dạng muối tiêu). Tất cả các mô hình được thực hiện với theo tác giãn nở cho hàm đầu ra trong quá trình nhớ lại. Các ảnh bị biến đổi có thể nhìn thấy trong Hình 5.1.



Hình 5.1: Thử nghiệm bộ nhớ tự liên kết với tập dữ liệu về số. Hàng đầu tiên chứa các ảnh huấn luyện gốc; Hàng thứ hai chứa các mẫu vào nhiễu bị biến đổi; Hàng thứ 3, 4, 5, 6 chứa mẫu ra từ mô hình của Junbo, Xiao, Sussner và Valle và ACAM.

Bảng 5.1: Kết quả của thử nghiệm bộ nhớ tự liên kết với tập dữ liệu về con số

	Junbo	Kosko	Xiao	IFAM	MAM	ACAM
Lỗi	0.434	0.914	0.418	0.434	0.434	0.346

Bảng 5.1 cho thấy tổng số lỗi của các mô hình khác nhau khi nhớ lại từ ảnh vào bị biến đổi trong chế độ tự liên kết. Số liệu từ bảng cho thấy ACAM có tổng lỗi ít nhất, trong khi mô hình của Xiao, IFAM của Sussner và Valle và MAM của Ritter có tổng số lỗi tương tự. Mô hình của Kosko có tổng lỗi cao nhất. Mô hình của Kosko thậm chí không thể cho ra kết quả hoàn hảo khi mẫu vào hoàn hảo trong nhiều trường hợp. Lý do các mô hình khác sản xuất tổng số lỗi lớn hơn so với mô hình ACAM là các mô hình này không thể làm việc tốt với cả hai loại nhiễu co rút và giãn nở trong khi ACAM có một cơ chế để giảm ảnh hưởng của nhiễu. Điều này có thể được nhìn thấy rõ ràng hơn trong Hình 5.1.

Bảng 5.2: Kết quả thử nghiệm của bộ nhớ liên kết khác loại với tập dữ liệu về con số

	Junbo	Kosko	Xiao	IFAM	MAM	ACAM
Lỗi	0.675	0.893	0.793	0.675	0.675	0.652

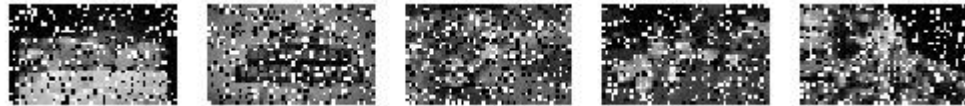
Trong chế độ liên kết khác loại, các cặp ảnh cần lưu trữ là ảnh của 0 và 1, 1 và 2, ... Bảng 5.2 cho thấy tổng lỗi của các mô hình khác nhau trong trường hợp này. Từ bảng cho thấy ACAM cũng tạo ra tổng lỗi ít nhất. Cần lưu ý rằng khi không có nhiễu hoặc chỉ có nhiễu co rút, mô hình của tác giả thực hiện hơi kém hơn IFAMs và MAMs vì cơ chế để giảm ảnh hưởng của nhiễu. Trong sự hiện diện của chỉ nhiễu giãn nở, mô hình của Xiao cũng thực hiện tốt hơn so với ACAM. Tuy nhiên, sự đánh đổi này là đáng giá để xem xét bởi vì trong thực tế mẫu vào hoàn hảo hay mẫu vào bị biến đổi chỉ bởi nhiễu co rút hay mẫu vào bị bóp méo chỉ bởi nhiễu giãn nở là không phổ biến.

5.5.2 Thử nghiệm với tập dữ liệu của Corel

Bộ dữ liệu này bao gồm các hình ảnh được lựa chọn từ cơ sở dữ liệu Corel (Hình 5.2). Các mẫu thử nghiệm được tạo ra từ các mẫu vào bằng cách tạo ra nhiễu muối tiêu ở mức 25% số lượng điểm ảnh. Hình 5.3 cho thấy một số mẫu nhiễu được tạo ra.



Hình 5.2: Một số ảnh từ tập dữ liệu của Corel dùng cho thử nghiệm



Hình 5.3: Các mẫu thử nghiệm được sinh ra từ các mẫu vào bằng nhiễu muối tiêu

Trong chế độ tự liên kết, 10 ảnh được sử dụng. Kết quả trong chế độ tự liên kết được trình bày trong bảng 5.3 cho thấy ACAM hiệu quả trong việc xử lý với nhiễu muối tiêu. Hình 5.4 cho thấy rõ rằng FAM cải tiến cải thiện mẫu ra nhiều hơn các mô hình khác.

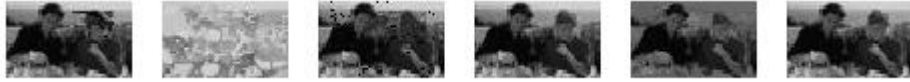
Chế độ liên kết khác loại được thử nghiệm với 10 cặp ảnh, trong đó các ảnh vào khác với ảnh ra. Như trong các thử nghiệm trước đó, mẫu vào bị biến đổi bởi nhiễu muối tiêu. Bảng 5.4 cũng cho thấy ACAM thực hiện tốt hơn so với các mô hình khác trong sự hiện diện của cả hai loại nhiễu co rút và giãn nở. Hình 5.5 cho kết quả so sánh FAM cải tiến với các mô hình khác.

Bảng 5.3: Kết quả của thử nghiệm bộ nhớ tự liên kết với tập dữ liệu của Corel

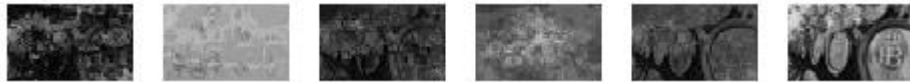
	Junbo	Kosko	Xiao	IFAM	MAM	ACAM
Lỗi	0.742	0.867	0.694	0.664	0.664	0.531

Bảng 5.4: Kết quả của thử nghiệm bộ nhớ liên kết khác loại với tập dữ liệu của Corel

	Junbo	Kosko	Xiao	IFAM	MAM	ACAM
Lỗi	0.795	1.018	0.702	0.624	0.624	0.548



Hình 5.4: Các mẫu từ tập dữ liệu của Corel được mô hình đưa ra phục hồi mẫu từ nhiễu muối tiêu tốt hơn các mô hình khác trong chế độ tự liên kết. Từ trái sang phải là các mẫu được phục hồi bởi mô hình của Junbo, Kosko, Xiao, Sussner và Valle, ACAM, và kết quả mong đợi.



Hình 5.5: Từ trái sang phải là các mẫu từ tập dữ liệu của Corel được phục hồi từ nhiễu muối tiêu trong chế độ liên kết khác loại bởi các mô hình của Junbo, Kosko, Xiao, Sussner và Valle, ACAM, và kết quả mong đợi.

3.6 Kết luận chương

Logic mờ cung cấp một thước đo lớn để giải quyết vấn đề không chắc chắn và không chính xác của ngôn ngữ. Hiện nay, nhiều hệ thống đang làm việc với các dữ liệu mờ, tri thức mờ như các mô tả trạng thái của bệnh nhân trong y học, định nghĩa các quy luật điều khiển hay quy luật thay đổi trong kinh tế, tài chính ... FAM là một ANN có ba ưu điểm quan trọng gồm chịu nhiễu, lưu trữ không giới hạn, hội tụ trong một lần lặp. Do đó, FAM được áp dụng cho nhiều bài toán thực như dự báo lượng nước của dòng sông, phục chế ảnh, dự đoán giá chứng khoán, thiết kế bộ điều khiển cho các thiết bị...

Tuy nhiên, các FAM đã công bố chỉ lưu trữ hiệu quả nội dung các mẫu hoặc chỉ lưu trữ tốt liên kết giữa các mẫu nên không có mô hình nào lưu trữ tốt cả nội dung và liên kết của các mẫu. Hơn nữa, các FAM trước đây chưa tập trung vào việc giảm ảnh hưởng của các mẫu vào nhiễu nên khả năng phục hồi các mẫu đã lưu trữ bị hạn chế khi mẫu vào có dạng nhiễu phức tạp. Do đó, việc phát triển một mô hình FAM lưu trữ tốt cả nội dung và liên kết của các mẫu và đồng thời giảm ảnh hưởng của các mẫu vào nhiễu là hết sức cần thiết cho các ứng dụng thực.

Một mô hình FAM lưu trữ cả nội dung các mẫu và sự liên kết giữa các cặp mẫu được đề xuất. Trong khi vẫn sở hữu các ưu điểm của FAM chuẩn, mô hình cải

tiến còn giảm ảnh hưởng của các mẫu vào nhiễu trong quá trình nhớ lại để chịu nhiễu tốt hơn với các cả hai dạng nhiễu co rút và giãn nở. Các thử nghiệm được làm trên các bộ dữ liệu khác nhau để chứng minh hiệu quả của FAM cải tiến. Kết quả thu được gợi ý rằng sự cải thiện trong việc học và lưu trữ cả nội dung và liên kết của các mẫu là có hiệu quả. Kết quả nghiên cứu này được công bố tại Tạp chí NeuroComputing với chỉ số SCIE (Công trình khoa học số 11), kỷ yếu có phản biện của Hội nghị quốc tế được xuất bản bởi Springer (Công trình khoa học số 5, 6, và 9).

KẾT LUẬN

Các cải tiến đề xuất cho từng mô hình trong luận án đã được trình bày và chứng minh bằng thực nghiệm với các kết quả sau:

Đề xuất một thuật toán học cho BAM để học nhanh và linh động hơn. Hơn nữa, BAM gắn với thuật toán học cải tiến còn lưu trữ và nhớ lại tốt với các cặp mẫu không trực giao. Năm thử nghiệm trong ứng dụng nhận dạng mẫu gồm nhận dạng vân tay, chữ viết tay, phương tiện giao thông, biển hiệu giao thông và tiền xu của Mỹ được làm để đánh giá khả năng nhớ lại của BAM cải tiến. Kết quả thực nghiệm cho thấy BAM đề xuất có khả năng nhớ lại tốt hơn các BAM khác trong chế độ tự liên kết.

Đưa ra hai luật học hiệu quả của Fuzzy ART để học tốt hơn mọi mẫu huấn luyện đồng thời giảm sự ảnh hưởng của các mẫu huấn luyện dị thường. Các tập dữ liệu chuẩn được chọn từ cơ sở dữ liệu UCI và Shape được dùng để đánh giá khả năng nhớ lại của Fuzzy ART cải tiến. Với luật học đề xuất thứ nhất, các thực nghiệm được làm trên 9 tập dữ liệu về các loài hoa lan, nguồn gốc rượu vang, các dạng vỡ của kính trong các vụ án hình sự và các hình dạng (bông hoa, bàn tay, pháo hoa,..) được tạo thành từ tập các điểm trong tập dữ liệu. Thực nghiệm thứ nhất cho thấy Fuzzy ART cải tiến phân cụm tốt hơn đáng kể so với các mô hình khác với các tập dữ liệu nhỏ, phức tạp. Với luật học cải tiến thứ hai, 7 tập dữ liệu về triệu chứng của bệnh ung thư, chất lượng rượu vang, các bài toán của Monk, các dữ liệu để thử nghiệm tâm lý học và các hình dạng được tạo thành từ tập các điểm trong tập dữ liệu được chọn. Kết quả của thực nghiệm thứ hai cho thấy khả năng phân cụm của Fuzzy ART cải tiến cao hơn đáng kể với các tập dữ liệu nhỏ, phức tạp. Đặc biệt, Fuzzy ART cải tiến phân cụm hiệu quả với các tập dữ liệu có số lượng mẫu cao.

Trình bày luật học cải tiến cho FAM để lưu trữ hiệu quả cả nội dung và sự liên kết giữa các cặp mẫu. Hơn nữa, FAM gắn với luật học cải tiến còn giảm sự ảnh hưởng của các mẫu vào nhiễu trong quá trình nhớ lại để chịu nhiễu tốt hơn với các

nhiều ở cả hai dạng co rút và giãn nở. Các thử nghiệm được làm trên các bộ ảnh về các con số và các ảnh được lấy ra từ cơ sở dữ liệu ảnh của hãng Corel trong các ứng dụng nhận dạng. Các FAM được thử nghiệm ở cả hai chế độ tự liên kết và liên kết khác loại. Kết quả thử nghiệm cho thấy FAM với luật học cải tiến có khả năng nhớ lại tốt hơn các FAM khác trong cả hai chế độ.

Với mong muốn đóng góp một phần vào việc nâng cao chất lượng của một số ANN nhớ thông tin, tác giả đã đề xuất một số cải tiến với ba mô hình ANN. Tuy nhiên, các kết quả nghiên cứu vẫn còn một số nhược điểm. Trong thời gian tới, tác giả mong muốn được tiếp tục phát triển các nghiên cứu về ANN để khắc phục các hạn chế và tăng tính hiệu quả của các ANN đã nghiên cứu.

DANH MỤC CÁC CÔNG TRÌNH KHOA HỌC CỦA TÁC GIẢ LIÊN QUAN ĐẾN LUẬN ÁN

1. Nong Thi Hoa, The Duy Bui (2012), **A new effective learning rule of Fuzzy ART**, *In Proceedings of 2012 Conference on Technologies and Applications of Artificial Intelligence*, IEEE Press, DOI 10.1109/TAAI.2012.60, pp 224-231.
2. Nong Thi Hoa, The Duy Bui (2012), **A new learning strategy for general BAMs**, *In Proceedings of 8th International Conference on Machine Learning and Data Mining*, Springer, LNAI 7376, pp 213-221.
3. Nong Thi Hoa, The Duy Bui (2012), **A fast effective learning strategy for bi-directional memory**, *In Proceedings of The first International Conference of Information Technology and Science*, SERSC, ISSN: 2287-1233/IST 2012, pp 183-185.
4. Nong Thi Hoa, The Duy Bui (2013), **A max-min learning rule of Fuzzy ART**, *In Proceedings of 10th IEEE RIVF International Conference on Computing and Communication*, IEEE Press, ISBN: 978-1-4799-1350-3/IEEE RIVF.2013, pp 53-57.
5. Nong Thi Hoa, The Duy Bui, Trung Kien Dang (2013), **Efficiency improvements for Fuzzy Associative Memory**, *In Proceedings of 10th International Symposium on Neural Network*, Springer, LNCS 7951, pp 36–43.
6. Pham Viet Binh, Nong Thi Hoa (2013), **Compare effective Fuzzy Associative Memories for grey-scale image recognition**, *In Proceedings of International Conference on Context - Aware Systems and Applications*, Springer, LNICST 109, pp 258-265.
7. Nong Thi Hoa, The Duy Bui (2013), **An improved learning algorithm of BAM**, *Journal of Science and Technology*, Thainguyen University, ISSN: 1859-2171, vol. 113(13), pp 61-65.
8. Nông Thi Hoa, Hoàng Trọng Vĩnh (2013), **Sử dụng mạng nơron nhân tạo Fuzzy ART để phân cụm dữ liệu**, *Tạp chí Khoa học và Công nghệ*, Đại học Thái Nguyên, ISSN: 1859-2171, tập 106(6), trang 49-54.
9. Pham Viet Binh, Nong Thi Hoa, Vu Duc Thai, Quach Xuan Truong (2014), **A new Fuzzy Associative Memory**, *In Proceedings of 2nd International Conference on Context - Aware Systems and Applications*, Springer, LNICST 128, pp 219-227
10. Nong Thi Hoa, The Duy Bui (2014), **An improved learning rule for Fuzzy ART**, *Journal of Information Science and Engineering*, 30(3), Institute of Information Science, Academia Sinica, pp 713-726. (ISI-indexed)
11. The Duy Bui, Nong Thi Hoa, Trung Kien Dang (2015), **Improving Learning Rule for Fuzzy Associative Memory with Combination of Content and Association**, *NeuroComputing*, Elsevier, 149(Part A), Elsevier, pp.59-64 (ISI-indexed).

TÀI LIỆU THAM KHẢO

1. Đặng Quang Á, Ứng dụng của mạng nơron trong tính toán, *Sách “Hệ mờ, mạng nơron và ứng dụng”*, Chủ biên: Bùi công Cường, Nguyễn Doãn Phước, Nhà XB Khoa học kỹ thuật, Hà nội, 2001, pp. 199-211.
2. M.E. Acevedo-mosqueda, C. Yáñez-márquez, I. López-yáñez (2006), “Alpha-Beta Bidirectional Associative Memories Based Translator”, *Journal of Computer Science*, vol. 6(5), pp. 190–194.
3. G.C. Anagnostopoulos, M. Georgiopoulos (2002), “Category regions as new geometrical concepts in Fuzzy-ART and Fuzzy-ARTMAP”, *Elsevier Science-Neural Network*, vol. 15, pp. 1205–1221.
4. P.V. Binh, N.T. Hoa, V.D. Thai, Q.X. Truong (2014), “A new Fuzzy Associative Memory”, *The first International Conference on Context - Aware Systems and Applications*, LNICST 128.
5. P.V. Binh, N.T. Hoa (2012), “Compare effective Fuzzy Associative Memories for grey-scale image recognition”, *The 2nd International Conference on Context - Aware Systems and Applications*, LNICST 109, pp. 258-265.
6. T. Burwick, F. Joublin (1998), “Optimal Algorithmic Complexity of Fuzzy ART”, *Kluwer Academic Publisher-Neural Processing Letters*, vol. 7, pp. 37–41.
7. M. Cano, Y. Dimitriadis, E. Gomez, J. Coronado (2001), “Learning from noisy information in FasArt and FasBack neuro-fuzzy systems”, *Elsevier Science-Neural Network*, vol. 14, pp. 407–425.
8. G.A. Carpenter, S. Grossberg, N. Markuron (1992), “Fuzzy ARTMAP-an addaptive resonance architecture for incremental learning of analog maps”, *The International Joint Conference on Neural Networks*, vol. 3.
9. G. Carpenter, S. Grossberg, D. B. Rosen (1991), “Fuzzy ART : Fast Stable Learning and Categorization of Analog Patterns by an Adaptive Resonance System”, *Pergamon Press-Neural network*, vol. 4, pp. 759–771.
10. S. Chartier, M. Boukadoum (2006), “A Bidirectional Heteroassociative Memory for Binary and Grey-Level Patterns”, *IEEE Transactions on Neural Networks*, vol. 17(2), pp. 385–396.
11. S. Chartier, M. Boukadoum, M. Amiri (2009), “BAM Learning of Nonlinearly Separable Tasks by Using an Asymmetrical Output Function and Reinforcement Learning”, *IEEE Transactions on Neural Networks*, vol. 20(8), pp. 1281–1292.
12. K. L. Chu, M. Ali, S.L. Wei (2013), “A Novel Complex-Valued Fuzzy ARTMAP for Sparse Dictionary Learning”, *Neural Information Processing*, LNCS, Vol 8226, pp. 360-368.

13. L.K. Chu, L.S. Wei, S. Manjeevan, L. Einly (2015), "Probabilistic ensemble Fuzzy ARTMAP optimization using hierarchical parallel genetic algorithms", *Neural Computing and Applications*, Vol. 26(2), pp. 263-276.
14. F. Chung, T. Lee (1994), "Towards a High Capacity Fuzzy Associative Memory Model", *IEEE International Conference on Neural Network*, vol. 3, pp. 1595–1599.
15. F. Chung, T. Lee (1996), "On fuzzy associative memories with multiple-rule storage capacity", *IEEE Transactions on Fuzzy System*, vol. 4(3). pp. 375–384.
16. M.Culloch, W. Pitts (1943), "A Logical Calculus of Ideas Immanent in Nervous Activity", *Bulletin of Mathematical Biophysics*, vol. 5(4), pp. 115–133.
17. I. Dagher, M. Georgiopoulos, G.L. Heileman, G. Bebis (1999), "An ordering algorithm for pattern presentation in fuzzy ARTMAP that tends to improve generalization performance.", *IEEE Transactions on Neural Networks*, vol. 10(4), pp. 768–78.
18. B.T. Duy, N.T. Hoa, D.T. Kien (2014), "Improving Learning Rule for Fuzzy Associative Memory with Combination of Content and Association", *NeuroComputing*, Elsevier, 149(Part A), Elsevier, pp.59-64 (ISI-indexed).
19. T. Eom, C. Choi, J. Lee (2002), "Generalized asymmetrical bidirectional associative memory for multiple association", *Applied Mathematics and Computation*, vol. 127(2–3), pp. 221–233.
20. T. Eom, C. Choi, J. Lee (1999), "Generalized Asymmetrical Bidirectional Associative Memory", *Machine Intelligence & Robotic Control*, vol. 1(1), pp. 43–45.
21. E.L. Esmi, P. Sussner (2010), "A Fuzzy Associative Memory Based on Kosko's Subsethood Measure", *The 2010 International Joint Conference on Neural Networks*, pp. 1–8.
22. Estivill-Castro, Vladimir (2002). "Why so many clustering algorithms — A Position Paper". *ACM SIGKDD Explorations Newsletter* vol. 4(1), pp. 65–75.
23. M. Georgiopoulos, H. Fernlund, G. Bebis, G. Heileman (1996), "FART and FARTMAP-Effects of the choice parameter", *Elsevier Science-Neural Network*, vol. 9, pp. 1541–1559.
24. S. Grossberg (1976), "Adaptive pattern classification and universal recoding, II: Feedback, expectation, olfaction and illusions", *Biological Cybernetics*, vol. 23, 187-212.
25. S. Grossberg (1980), "How does a brain build a cognitive code", *Studies of mind and brain: Neural principles of learning, perception, development, cognition, and motor control*, Boston, MA: Reidel Press.

26. D.O. Hebb (1949), "*Organization of Behavior: a Neuropsychological Theory*", New York, John Wiley.
27. N.T. Hoa, B.T. Duy (2014), "An improved learning rule of Fuzzy ART", *Journal of Information Science and Engineering*, 30(3), pp. 713-726, Institute of Information Science, Academia Sinica.
28. N.T. Hoa, B.T. Duy (2013), "A max-min learning rule of Fuzzy ART", *The 10th IEEE RIVF International Conference on Computing and Communication Technologies*, pp. 53-57.
29. N.T. Hoa, B.T. Duy Bui, D.T. Kien (2013), "Efficiency improvements for Fuzzy Associative Memory", *The 10th International Symposium on Neural Network*, pp. 36–43.
30. N.T. Hoa, B.T. Duy (2012), "A new effective learning rule of Fuzzy ART", *The 2012 Conference on Technologies and Applications of Artificial Intelligence*, pp. 224-231.
31. N.T. Hoa, B.T. Duy (2012), "A new learning strategy for general BAMs", *The 8th International Conference on Machine Learning and Data Mining*, pp. 213-221
32. N.T. Hoa, B.T. Duy (2013), "An improved learning algorithm of BAM", *Journal of Science and Technology*, Thainguyen University, vol. 113(13), pp. 61-65.
33. N.T. Hoa, B.T. Duy (2012), "A fast effective learning strategy for bi-directional memory", *The first International Conference of Information Technology and Science*, SERSC Press, ISSN: 2287-1233/IST 2012, pp 183-185.
34. N.T. Hoa, H.T. Vinh (2013), "Sử dụng mạng nơron nhân tạo Fuzzy ART để phân cụm dữ liệu", *Tạp chí Khoa học và Công nghệ, Đại học Thái Nguyên*, ISSN: 1859-2171, tập 106(6) năm 2013, trang 49-54.
35. J.J.Hopfield (1982), "Neural networks and physical systems with emergent collective computational abilities", *The National Academy of Sciences of the USA*, vol. 79(8), pp. 2554–2558.
36. M.H. Hassoun (1993), "Dynamic associative neural memories", *Associative Neural Memories: Theory and Implementation*, Oxford University Press, Oxford, U.K.
37. J. Huang, M. Georgiopoulos, G.L. Heileman (1995), "Fuzzy ART Properties", *Elsevier Science-Neural Network*, vol. 8(2), pp. 203–213.
38. H. Isawa, H. Matsushita, Y. Nishio (2008), "Improved Fuzzy Adaptive Resonance Theory Combining Overlapped Category in Consideration of Connections", *IEEE Workshop on Nonlinear Circuit Networks*, pp. 8–11.

39. H. Isawa, M. Tomita, H. Matsushita, Y. Nishio (2007), "Fuzzy Adaptive Resonance Theory with Group Learning and its Applications", *International Symposium on Nonlinear Theory and its Applications*, vol. 1, pp. 292–295.
40. F. Junbo, J. Fan, S. Yan (1994), "A learning rule for Fuzzy Associative Memory", *IEEE 8th International Conference on Neural Networks*, vol. 7, pp. 4273–4277.
41. R. Kenaya, K.C. Cheok (2008), "Euclidean ART Neural Networks", *The World Congress on Engineering and Computer Science*, Vol. 1.
42. K. Kobayashi, S. Mizuno, T. Kuremoto, M. Obayashi (2005), "A Reinforcement Learning System Based on State Space Construction Using Fuzzy ART", *SICE Annual Conference*, vol. 1, pp. 3653–3658.
43. S.G. Kong, B. Kosko (1992), "Adaptive fuzzy systems for backing up a truck-and-trailer", *IEEE Transactions on Neural Networks*, vol.3, pp. 211–223.
44. B. Kosko (1992), *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*, Prentice Hall, Englewood Cliffs, New Jersey.
45. B. Kosko (1988), "Bidirectional Associative Memory", *IEEE Transaction on Systems, Man, and Cybernetic*, vol. 18(1), pp. 49–60.
46. C.S. Leung (1994), "Optimum Learning for Bidirectional Associative Memory in the Sense of Capacity", *IEEE Transactions on Systems, Man and Cybernetics*, vol. 24 (5), pp. 791–796.
47. L. Li, J. Yang, W. Wu, T. Wu (2009), "An intuitionistic fuzzy associative memory network and its learning rule", *IEEE International Conference on Granular Computing*, pp. 350–353.
48. C. Lin, C. Lin, C.S. G. Lee (1995), "Fuzzy adaptive learning control network with on-line neural learning", *Elsevier Science-Fuzzy sets and Systems*, vol. 71, pp. 25–45.
49. B. Moore (1989), "ART 1 and pattern clustering", *The Connectionist Models Summer School*, in D. Touretzky, G. Hinton, & T. Sejnowski (Eds.) San Mateo, CA: Morgan Kaufmann Publishers, pp. 174-185.
50. W. Pedrycz (1993), "Fuzzy neural networks and neurocomputations", *Fuzzy Sets and Systems*, vol. 56, pp. 1–28.
51. J.B. Queen (1967), "Some methods for classification and analysis of multivariate observations", *The 5th Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, pp. 281–297.
52. G. Ritter, P. Sussner, J.D. Leon (1998), "Morphological associative memories", *IEEE Transactions on Neural Networks*, vol. 9, pp. 281-293.
53. F. Rosenblatt (1958), "The perceptron: A probabilistic model for information storage and organization in the brain", *Psychological review*, vol. 65, pp. 386-408.

54. D. Shen, J.B. Cruz (2003), "Encoding strategy for maximum noise tolerance Bidirectional Associative Memory", *IEEE Transactions on Neural Networks*, vol. 2(16), pp. 293–300.
55. W.Y. Shen, Y.S. Keem, Y.J. Hwa, T.C. Shing (2014), "A Truly Online Learning Algorithm using Hybrid Fuzzy ARTMAP and Online Extreme Learning Machine for Pattern Classification", *Neural Processing Letters*, doi 10.1007/s11063-014-9374-5.
56. J. Serra (1982), *Image Analysis and Mathematical Morphology*, Academic Press, London.
57. H. Shi, Y. Zhao, X. Zhuang (1998), "A general model for bidirectional associative memories", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 28(4), pp. 511–519.
58. P. Sussner, M.E. Valle (2006), "Implicative Fuzzy Associative Memories", *IEEE Transactions on Fuzzy System*, vol. 14(6), pp. 793–807.
59. P. Sussner, M.E. Valle (2008), *Handbook of Granular Computing*, Willey, pp. 1–41.
60. Y. Takefuji (1992), *Neural Network Parallel Computing*, Kluwer Acad. Publish.
61. A.H. Tan (1995), "Adaptive Resonance Associative Map", *Elsevier Science-Neural Network*, vol. 8(3), pp. 437–446.
62. M.E. Valle (2010), "A New Class of Implicative FAM for the Reconstruction of Gray-Scale Images Corrupted by Salt and Pepper Noise", *The 7th Brazilian Symposium on Neural Networks*, pp. 200–205.
63. V. Vidya, T. R. Indhu, V. K. Bhadrán, R. Ravindra Kumar (2013), "Malayalam Offline Handwritten Recognition Using Probabilistic Simplified Fuzzy ARTMAP", *Intelligent Informatics, Advances in Intelligent Systems and Computing*, Volume 182, pp 273-283.
64. R.A. Vázquez, H. Sossa, B.A. Garro, C.D México (2006), "A New Bi-directional Associative Memory", *The 5th Mexican International Conference on Artificial Intelligence*, pp. 367 – 380.
65. S.T. Wang, H.J. Lu (2004), "On New Fuzzy Morphological Associative Memories", *IEEE International Conference on Systems, Man, and Cybernet*, vol. 12(3), pp. 316–323.
66. T. Wang, X. Zhuang (1992), "Weighted Learning of Bidirectional Associative Memories by Global Minimization", *IEEE Transaction on Neural Networks*, vol. 3(6), pp. 1010–1018.
67. T. Wang, X. Zhuang, X. Xing (1994), "Memories with Optimal Stability", *IEEE Transactions on Neural Networks*, vol. 24(5), pp. 778–790.

68. Y.F. Wang, J.R. Cruz, J.R. Mulligan (1990a), “Two coding strategies for bidirectional associative memory.”, *IEEE Transactions on Neural Networks*, vol. 1(1), pp. 81–92.
69. Y.F. Wang, J.R. Cruz, J.R. Mulligan (1990b), “On multiple training for bidirectional associative memory.”, *IEEE Transactions on Neural Networks*, vol. 1(3), pp. 275–276.
70. Y.F. Wang, J.R. Cruz, J.R. Mulligan (1991), “Guaranteed recall of all training pairs for BAM”, *IEEE Transactions on Neural Networks*, vol. 2(6), pp. 559–566.
71. P. Xiao, F. Yang, Y. Yu (1997), “Max-Min Encoding Learning Algorithm for Fuzzy Max-Multiplication Associative Memory Networks”, *IEEE International Conference on Systems, Man, and Cybernet*, vol. 4, pp. 3674–3679.
72. Z. Xu, X. He (1994), “Asymmetric Bidirectional Associative Memories”, *IEEE International Conference on Systems, Man, and Cybernet*, vol. 24(10), pp. 729–735.
73. A. Yousuf, Y.L. Murphey (2010), “A Supervised Fuzzy Adaptive Resonance Theory with Distributed Weight Update”, *The 7th International Symposium on Neural Network*, Springer, vol. Part I, LNCS, no. 6063, pp. 430–435.
74. L.A. Zadeh (1979), “Fuzzy sets and information granularity”. *Advances in Fuzzy Set Theory and Applications*, North Holland, Amsterdam, pp. 3–18.
75. S. Zeng, W. Xu, J. Yang (2008), “Research on Properties of Max-Product Fuzzy Associative Memory Networks”, *The 8th International Conference on Intelligent Systems Design and Applications*, pp. 438–443.
76. X. Zhuang, Y. Huang, S.S. Chen (1993), “Better learning for bidirectional associative memory”, *Neural Networks*, vol. 6(8), pp. 1131–1146.