

ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ

NGUYỄN THỊ QUỲNH HOA

NGHIÊN CỨU VÀ PHÁT TRIỂN GIẢI PHÁP
THEO DÕI CHUYỂN ĐỘNG MẮT
HỖ TRỢ TRẺ RỐI LOẠN PHÁT TRIỂN

LUẬN ÁN TIẾN SĨ KHOA HỌC MÁY TÍNH

Hà Nội - 2026

ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ

NGUYỄN THỊ QUỲNH HOA

NGHIÊN CỨU VÀ PHÁT TRIỂN GIẢI PHÁP
THEO DÕI CHUYỂN ĐỘNG MẮT
HỖ TRỢ TRẺ RỐI LOẠN PHÁT TRIỂN

Chuyên ngành Khoa học máy tính

Mã số: 9480101

LUẬN ÁN TIẾN SĨ KHOA HỌC MÁY TÍNH

NGƯỜI HƯỚNG DẪN KHOA HỌC:

- PGS.TS. Lê Thanh Hà
- TS. Ngô Thị Duyên

Hà Nội - 2026

LỜI CAM ĐOAN

Tôi xin cam đoan luận án **“Nghiên cứu và phát triển giải pháp theo dõi chuyển động mắt hỗ trợ trẻ rối loạn phát triển”** là công trình nghiên cứu do chính tôi thực hiện. Các số liệu và kết quả được trình bày trong luận án là hoàn toàn trung thực, khách quan và chưa từng được công bố trong bất kỳ công trình khoa học của cá nhân nào khác.

- Tôi đã trích dẫn đầy đủ các tài liệu tham khảo, công trình nghiên cứu liên quan ở trong nước và quốc tế. Ngoại trừ các tài liệu tham khảo này, luận án hoàn toàn là công việc của riêng tôi.
- Trong các công trình khoa học được công bố trong luận án, tôi đã thể hiện rõ ràng và chính xác đóng góp của các đồng tác giả và những gì do tôi đã đóng góp.
- Luận án được hoàn thành trong thời gian tôi làm nghiên cứu sinh tại phòng Thí nghiệm Tương tác người-máy tại Khoa Công nghệ thông tin - Trường Đại học Công nghệ, Đại học Quốc gia Hà Nội.

Tác giả:

Hà Nội:

LỜI CẢM ƠN

Lời đầu tiên, tôi xin bày tỏ lòng biết ơn sâu sắc nhất tới PGS.TS. Lê Thanh Hà và TS. Ngô Thị Duyên, tập thể giảng viên hướng dẫn, thầy cô đã luôn tận tâm chỉ bảo và dành nhiều tâm huyết để định hướng cho tôi trong suốt quá trình thực hiện nghiên cứu và hoàn thành luận án. Một vinh dự lớn cho tôi được học tập, nghiên cứu dưới sự hướng dẫn của các thầy, cô. Sự nghiêm túc trong khoa học cùng những định hướng chuyên môn quý báu của thầy cô chính là kim chỉ nam giúp tôi vượt qua những thách thức để hoàn thiện luận án tại trường Đại học Công nghệ, Đại học Quốc gia Hà Nội.

Tôi xin chân thành cảm ơn Ban giám hiệu, tập thể thầy/cô cùng cán bộ tại trường Đại học Công nghệ, Đại học Quốc gia Hà Nội đã tạo điều kiện thuận lợi về môi trường học thuật cũng như cơ sở vật chất để tôi thực hiện các nghiên cứu của luận án. Đặc biệt, tôi xin gửi lời tri ân tới các em học sinh, các bậc phụ huynh và các trung tâm giáo dục chuyên biệt đã tin tưởng cộng tác trong quá trình thực hiện luận án. Tôi xin được bày tỏ sự biết ơn sâu sắc đến các thầy cô trong phòng phòng Thí nghiệm Tương tác người-máy tại Khoa Công nghệ thông tin, Trường Đại học Công nghệ, Đại học Quốc gia Hà Nội vì sự giúp đỡ của các thầy, cô về các đóng góp rất hữu ích cho luận án.

Tôi xin trân trọng cảm ơn Khoa Công nghệ thông tin, Phòng Đào tạo và các phòng ban Trường Đại học Công nghệ đã tạo điều kiện thuận lợi cho tôi trong suốt quá trình thực hiện luận án.

Tôi cũng bày tỏ sự biết ơn đến trường Đại học Sư phạm Hà Nội đã tạo điều kiện về thời gian và tài chính cho tôi thực hiện luận án này. Tôi muốn cảm ơn đến Ban Chủ nhiệm khoa, các cán bộ giảng viên khoa Công nghệ thông tin, Trường Đại học Sư phạm Hà Nội đã cổ vũ, động viên và tạo điều kiện về thời gian cho tôi trong suốt quá trình nghiên cứu.

Tôi muốn cảm ơn đến tất cả những người bạn, các anh chị em nghiên cứu sinh và những người đồng nghiệp của tôi. Những người đã luôn chia sẻ với tôi những khó khăn và động viên giúp đỡ tôi khi tôi cần và tôi luôn ghi nhớ điều đó.

Tôi cũng xin gửi lời cảm ơn tới các cộng sự và các chuyên gia phản biện quốc tế đã giúp tôi hoàn thiện các công bố khoa học trong quá trình

làm nghiên cứu sinh. Cuối cùng, tôi xin dành lời cảm ơn chân thành nhất tới gia đình và bạn bè, những người luôn là điểm tựa tinh thần vững chắc, tiếp thêm sức mạnh để tôi kiên định hoàn thành chặng đường nghiên cứu này.

Mặc dù đã có nhiều cố gắng, luận án không tránh khỏi những thiếu sót. Tôi rất mong nhận được những ý kiến đóng góp quý báu để công trình được hoàn thiện hơn.

NCS. Nguyễn Thị Quỳnh Hoa

TÓM TẮT

Công nghệ theo dõi chuyển động mắt là một giải pháp đo lường khách quan trong các nghiên cứu về hành vi của con người. Thông qua cơ chế đo lường không xâm lấn, công nghệ này có khả năng ghi nhận tọa độ điểm nhìn theo thời gian thực, từ đó cung cấp các chỉ số về quỹ đạo chuyển động của mắt trong quá trình tiếp nhận thông tin thị giác. Hiện nay, giải pháp này đang được ứng dụng trong nhiều lĩnh vực như y tế, thương mại và giáo dục. Đặc biệt, trong bối cảnh hỗ trợ nhóm trẻ rối loạn phát triển thần kinh, công nghệ theo dõi chuyển động mắt cung cấp cơ sở định lượng để phân tích trạng thái nhận thức, phản ánh khách quan mức độ chú ý thị giác và năng lực xử lý thông tin trực quan. Trên cơ sở đó, nguồn dữ liệu này được ứng dụng nhằm hỗ trợ phát hiện sớm các dấu hiệu phát triển không điển hình và xây dựng lộ trình can thiệp giáo dục phù hợp với đặc điểm của từng trẻ.

Rối loạn phát triển thần kinh là nhóm tình trạng xuất hiện sớm từ thời thơ ấu, tác động trực tiếp đến sinh hoạt và năng lực học tập. Trong đó, ba dạng tiêu biểu bao gồm: Rối loạn phổ tự kỷ đặc trưng bởi những khiếm khuyết trong giao tiếp xã hội và hành vi lặp khuôn; chứng khó đọc đặc trưng bởi những rào cản trong nhận diện từ vựng và giải mã ngôn ngữ; và rối loạn tính toán đặc trưng bởi sự hạn chế trong tiếp cận khái niệm số lượng và xử lý quy luật logic. Sự đa dạng và khác biệt trong năng lực tư duy cũng như biểu hiện hành vi của từng trẻ đòi hỏi các phương pháp can thiệp phải mang tính cá nhân hóa. Để đáp ứng yêu cầu này, công nghệ theo dõi chuyển động mắt được ứng dụng như một công cụ đo lường và hỗ trợ hiệu quả.

Dù có tiềm năng, việc ứng dụng công nghệ theo dõi chuyển động mắt trong hỗ trợ nhóm trẻ rối loạn phát triển vẫn gặp một số rào cản. Trước hết, các bộ dữ liệu hiện tại còn hạn chế về cấu trúc và quy mô, gây khó khăn cho việc xây dựng các mô hình nhận diện hành vi nhận thức phức tạp. Thách thức này biểu hiện qua ba khía cạnh. Thứ nhất là sự thiếu hụt các bộ dữ liệu đa phương thức có tính đại diện cho bối cảnh thực tế. Các nghiên cứu hiện tại chủ yếu thu thập dữ liệu tọa độ điểm nhìn đơn lẻ đối với các kích thích thị giác dạng tĩnh, thiếu sự đồng bộ với tín hiệu âm thanh và thao tác tương tác. Phương pháp thu thập này khiến các bộ dữ liệu chưa phản ánh đầy đủ quá trình nhận thức chủ động của trẻ trong môi trường học tập tự nhiên, từ đó gây khó khăn cho việc xác lập mối liên hệ giữa đặc trưng hành

vi thị giác và năng lực tư duy. Thứ hai là những giới hạn kỹ thuật trong quá trình trích xuất và xử lý dữ liệu chuyển động mắt. Về mặt không gian, việc áp dụng các phương pháp xác định vùng quan tâm thủ công hoặc tĩnh đối với chuỗi khung hình động thường phát sinh sai số đáng kể. Về mặt quy mô, sự hạn chế và tình trạng phân bố mắt cân bằng của dữ liệu y sinh làm suy giảm hiệu năng huấn luyện của các mô hình học máy. Việc khuyết thiếu các giải pháp tăng cường dữ liệu có khả năng duy trì đặc tính sinh lý vi mô của nhãn cầu là một rào cản kỹ thuật cần giải quyết. Thứ ba là những hạn chế trong các cơ chế hỗ trợ can thiệp giáo dục hiện tại. Do thiếu các công cụ phân tích dữ liệu hành vi theo thời gian thực, việc điều chỉnh lộ trình can thiệp hiện nay vẫn chủ yếu phụ thuộc vào các đánh giá định tính từ chuyên gia hoặc chỉ thay đổi độ khó dựa trên kết quả hoàn thành nhiệm vụ. Sự khuyết thiếu cơ chế phản hồi dựa trên trạng thái nhận thức tức thời làm giảm khả năng tự động hóa và cá nhân hóa lộ trình can thiệp đối với từng trẻ.

Nhằm khắc phục các thách thức nêu trên, luận án thực hiện nghiên cứu và đề xuất các giải pháp công nghệ, trong đó trọng tâm là công nghệ theo dõi chuyển động mắt để hỗ trợ giáo dục và can thiệp cho trẻ rối loạn phát triển. Mục tiêu này được cụ thể hóa thông qua ba hướng nghiên cứu chính, mỗi hướng giải quyết một nhiệm vụ tương ứng trong quy trình hỗ trợ trẻ. Thứ nhất, luận án đề xuất phương pháp xây dựng các bộ dữ liệu đa phương thức cho nhóm trẻ rối loạn phổ tự kỉ và nhóm trẻ rối loạn tính toán. Bộ dữ liệu này tích hợp thông tin chuyển động mắt cùng các hành vi tương tác gồm giọng nói và dữ liệu tương tác với phần mềm, tạo cơ sở thông tin đầu vào cho các thuật toán học máy nhằm phân loại đặc trưng hành vi và đánh giá trạng thái nhận thức ở các giai đoạn sau. Thứ hai, luận án đề xuất các phương pháp xử lý và nâng cao năng lực khai thác dữ liệu chuyển động mắt. Cụ thể, một kiến trúc lai tích hợp mô hình phân đoạn hình ảnh và thuật toán luồng quang học được phát triển nhằm tự động hóa quy trình trích xuất đặc trưng thị giác trên kích thích động. Song song đó, để giải quyết rào cản thiếu hụt dữ liệu, luận án đề xuất phương pháp tăng cường quỹ đạo ánh nhìn dựa trên thuật toán tối ưu tiến hóa và mô hình chuyển động Brownian phân số (fBm). Giải pháp này hỗ trợ sinh dữ liệu tăng cường có khả năng duy trì cấu trúc phổ nhiễu hồng $1/f$ trong chuyển động của nhãn cầu, qua đó cải thiện hiệu năng phân loại của các mô hình học máy ở giai đoạn sau. Thứ ba, luận án kết hợp công nghệ theo dõi chuyển động mắt cùng các giải pháp kỹ thuật liên quan để đề xuất các phương pháp hỗ trợ

can thiệp cá nhân hóa cho trẻ rối loạn phát triển thần kinh.

Kết quả thực nghiệm cho thấy tính khả thi và khả năng ứng dụng của các phương pháp được đề xuất trong môi trường giáo dục và can thiệp thực tế. Giá trị khoa học của những kết quả này đã được ghi nhận thông qua các công trình khoa học được liệt kê tại mục “DANH MỤC CÁC CÔNG TRÌNH KHOA HỌC CỦA TÁC GIẢ LIÊN QUAN TỚI LUẬN ÁN”.

Từ khóa: Theo dõi chuyển động mắt (Eye-tracking), Tương tác người-máy (Human Computer Interaction), Rối loạn phát triển (Developmental Disorders), Rối loạn phổ tự kỉ (Autism Spectrum Disorder – ASD), Rối loạn tính toán (Dyscalculia), Chứng khó đọc (Dyslexia), Can thiệp (Intervention)

Mục lục

Lời cam đoan	i
Lời cảm ơn	ii
Mục lục	vii
Danh mục các từ viết tắt	xiii
Danh mục bảng	xvi
Danh mục hình vẽ	xvii
MỞ ĐẦU	1
1 Đặt vấn đề	1
2 Mục tiêu nghiên cứu	4
3 Đóng góp của luận án	4
4 Cấu trúc luận án	6
Chương 1. Tổng quan về công nghệ theo dõi chuyển động mắt hỗ trợ trẻ rối loạn phát triển	8
1.1 Công nghệ theo dõi chuyển động mắt	8
1.1.1 Các đặc trưng của dữ liệu chuyển động mắt	10
1.1.2 Các phương pháp trực quan hóa dữ liệu chuyển động mắt	11
1.1.3 Phương pháp xác định vùng quan tâm trong phân tích dữ liệu chuyển động mắt	12
1.2 Đặc trưng hành vi thị giác của trẻ rối loạn phát triển	13
1.2.1 Rối loạn phổ tự kỉ	13
1.2.2 Rối loạn tính toán	14
1.2.3 Chứng khó đọc	15
1.3 Hệ thống hóa giải pháp theo dõi chuyển động mắt hỗ trợ trẻ rối loạn phát triển	16
1.3.1 Phương pháp luận	17

1.3.2	Công nghệ theo dõi chuyển động mắt trong phân tích đặc trưng thị giác	19
1.3.3	Công nghệ theo dõi chuyển động mắt trong hỗ trợ chẩn đoán	20
1.3.4	Công nghệ theo dõi chuyển động mắt trong hỗ trợ can thiệp và đánh giá hiệu quả can thiệp	21
1.3.5	Thảo luận	23
1.3.6	Các khoảng trống nghiên cứu	24
1.3.7	Giới hạn phạm vi nghiên cứu	25
1.4	Các bộ dữ liệu tham chiếu và tiêu chuẩn đánh giá	26
1.4.1	Các bộ dữ liệu tham chiếu công khai	26
1.4.2	Hệ thống các độ đo đánh giá	27
	Kết chương	27
 Chương 2. Xây dựng bộ dữ liệu đa phương thức của trẻ rối loạn phát triển 29		
2.1	Giới thiệu	29
2.1.1	Vai trò của dữ liệu chuyển động mắt trong đánh giá trẻ rối loạn phát triển	29
2.1.2	Động lực và đóng góp	30
2.2	Quy trình xây dựng bộ dữ liệu đa phương thức	31
2.2.1	Quy trình tổng quan	31
2.2.2	Nguyên tắc chọn mẫu và đạo đức nghiên cứu	32
2.3	Xây dựng bộ dữ liệu hành vi thị giác của trẻ rối loạn phổ tự kỉ	33
2.3.1	Thiết kế kịch bản kích thích thị giác	33
2.3.2	Thực hiện thu dữ liệu	35
2.3.2.1	Đối tượng tham gia	35
2.3.2.2	Quy trình triển khai	36
2.3.3	Tiền xử lý và đồng bộ hóa dữ liệu	37
2.3.4	Tổ chức lưu trữ dữ liệu	38
2.4	Xây dựng bộ dữ liệu đa phương thức của trẻ rối loạn tính toán	39
2.4.1	Thiết kế kịch bản học liệu tương tác	39
2.4.2	Thực hiện thu nhận đa luồng tín hiệu	40
2.4.2.1	Đối tượng tham gia	40
2.4.2.2	Hệ thống thu nhận	41
2.4.3	Đồng bộ và hợp nhất dữ liệu đa phương thức	42
2.4.4	Tổ chức lưu trữ dữ liệu	44

2.5	Thảo luận về cơ sở khoa học và kiểm soát rủi ro	44
2.5.1	Quy mô mẫu và năng lực khái quát hóa của dữ liệu sinh trắc học	45
2.5.2	Đặc tính phân phối thừa của dữ liệu hướng sự kiện . .	45
2.5.3	Đặc điểm của dữ liệu đặc trưng đa phương thức	46
2.5.4	Kiểm soát dữ liệu hao hụt và rủi ro thiên lệch	46
2.6	Đánh giá độ tin cậy của bộ dữ liệu	47
	Kết chương	47
 Chương 3. Phương pháp trích xuất đặc trưng thị giác và tăng cường dữ liệu chuyển động mắt 49		
3.1	Giới thiệu	50
3.1.1	Vai trò và rào cản trong phân tích dữ liệu chuyển động mắt	50
3.1.2	Động lực và đóng góp	50
3.2	Phương pháp lai trích xuất đặc trưng thị giác tự động . . .	51
3.2.1	Bước 1: Phát hiện điểm cố định bằng thuật toán PeyeMMV	52
3.2.2	Bước 2: Xác định và theo dõi vùng quan tâm động . .	54
3.2.2.1	Cơ chế phân đoạn và lan truyền không gian và thời gian	54
3.2.2.2	Thuật toán tổng thể phương pháp lai trích xuất đặc trưng	56
3.2.3	Bước 3: Ánh xạ ngữ nghĩa và trích xuất đặc trưng chiến lược	56
3.3	Thực nghiệm và đánh giá phương pháp trích xuất đặc trưng thị giác tự động	58
3.3.1	Đánh giá kỹ thuật nhận diện vùng quan tâm tự động .	58
3.3.1.1	Hiệu năng phân đoạn vùng quan tâm	58
3.3.1.2	Cơ chế hợp nhất dữ liệu và tối ưu hóa tính toán	60
3.3.1.3	Đánh giá độ trễ và khả năng đáp ứng thời gian thực	61
3.3.2	Thiết lập thực nghiệm phân tích và phân loại lâm sàng	61
3.3.3	Đánh giá đặc trưng thị giác và hiệu năng hỗ trợ chẩn đoán ở trẻ rối loạn phổ tự kỉ	63
3.3.3.1	Phân tích thống kê và trực quan hóa hành vi thị giác	64

3.3.3.2	Thiết lập luồng trích xuất đặc trưng và tối ưu hóa mô hình phân loại	67
3.3.3.3	Kiểm chứng hiệu năng hỗ trợ phân loại và xác định dấu ấn sinh học	69
3.3.3.4	Đánh giá đối chứng hiệu quả của phương pháp trích chọn đặc trưng	70
3.3.3.5	Đối sánh với các nghiên cứu liên quan	74
3.3.4	Đánh giá đặc trưng thị giác và cơ chế nhận thức ở trẻ rối loạn tính toán	75
3.3.4.1	Phân tích thời lượng quan sát và ma trận chuyển đổi ánh nhìn	75
3.3.4.2	Phân tích đặc trưng nhận thức	77
3.3.4.3	Phân tích kết hợp nhiều nguồn dữ liệu và mối liên hệ hành vi	80
3.4	Phương pháp tăng cường dữ liệu chuyển động mắt dựa trên tối ưu tiến hóa	80
3.4.1	Bài toán tăng cường dữ liệu chuyển động mắt	80
3.4.2	Mô hình hóa động lực học bằng chuyển động Brownian phân số (fBm)	81
3.4.2.1	Đặc tính sinh lý của các vi chuyển động mắt	81
3.4.2.2	Mô hình hóa đặc tính động học nhân cầu dựa trên fBm	81
3.4.2.3	Kiến trúc tổng thể của hệ thống	83
3.4.2.4	Cấu trúc hàm đánh giá dựa trên thuật toán PeyMMV	83
3.4.2.5	Thiết kế hàm thích nghi	83
3.5	Thực nghiệm và đánh giá hiệu quả của phương pháp tăng cường dữ liệu	84
3.5.1	Đánh giá tỉ lệ phát hiện và độ phân tán không gian	84
3.5.2	Đánh giá độ chân thực phổ và đặc tính nhiễu hồng	85
3.5.3	Khả năng bảo tồn các đặc trưng nhận thức	85
3.5.4	Đánh giá hiệu năng phân loại chẩn đoán	86
	Kết chương	87
Chương 4.	Các phương pháp can thiệp cá nhân hóa hỗ trợ trẻ rối loạn phát triển	90
4.1	Giới thiệu	90

4.1.1	Bối cảnh và định hướng phương pháp luận	90
4.1.2	Động lực và đóng góp	91
4.2	Kiến trúc hệ thống can thiệp vòng kín	92
4.2.1	Mô hình khái niệm và vai trò của dữ liệu chuyển động mắt	92
4.2.2	Cơ sở phương pháp thực nghiệm: Phương pháp đánh giá đa cấp độ và thiết kế thực nghiệm trường hợp đơn	93
4.2.3	Cơ sở lựa chọn mô hình tính toán: Hệ cơ sở luật và mạng Bayes động	94
4.3	Hệ thống can thiệp cá nhân hóa cho trẻ rối loạn phổ tự kỉ dựa trên cơ sở luật suy diễn và dữ liệu thị giác	95
4.3.1	Cấu trúc chu trình can thiệp năm giai đoạn tích hợp công nghệ theo dõi chuyển động mắt	95
4.3.2	Mô hình hóa hệ cơ sở luật suy diễn cá nhân hóa	96
4.3.2.1	Cấu trúc hệ thống luật suy luận trạng thái hành vi	98
4.3.2.2	Mô hình ánh xạ đề xuất hiệu chỉnh và tối ưu hóa can thiệp	99
4.3.3	Phương pháp đánh giá hiệu quả can thiệp thông qua thiết kế thực nghiệm trường hợp đơn và thống kê Tau-U	100
4.3.3.1	Đặc tả cấu trúc và ý nghĩa của dữ liệu thô đầu vào	102
4.3.3.2	Quy trình trích xuất và thiết lập thang đo nhận thức	102
4.3.3.3	Mô hình giả thuyết và phương pháp kiểm định Tau-U	103
4.3.3.4	Phân tích kết quả thực nghiệm trường hợp đơn	103
4.4	Các nền tảng số hỗ trợ can thiệp trẻ rối loạn tính toán và trẻ mắc chứng khó đọc	106
4.4.1	Nền tảng web EzMath hỗ trợ trẻ rối loạn tính toán	106
4.4.2	Nền tảng di động EzRead hỗ trợ trẻ mắc chứng khó đọc	111
4.5	Mô hình suy diễn trạng thái nhận thức từ dữ liệu đa phương thức	116
4.5.1	Kiến trúc vòng lặp kép dựa trên đồ thị phụ thuộc phân tích nhiệm vụ	116
4.5.2	Mạng Bayes động và mô hình hóa nhận thức	121
4.5.2.1	Các biến trạng thái và tham số xác suất	121

4.5.2.2	Thuật toán cập nhật suy diễn Bayes	122
4.5.2.3	Hạn chế về hiện tượng bất định danh tham số	124
4.5.2.4	Mạng Bayes động cải tiến tích hợp hàm xác suất quan sát động	125
4.5.2.5	Phân tích độ phức tạp và tính hội tụ	126
4.6	Thảo luận chung	127
	Kết chương	127
KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN		130
DANH MỤC CÔNG TRÌNH KHOA HỌC		135
TÀI LIỆU THAM KHẢO		136

DANH MỤC CÁC TỪ VIẾT TẮT

Từ viết tắt	Tiếng Anh	Tiếng Việt
ABLLS-R	Assessment of Basic Language and Learning Skills, Revised	Đánh giá các kỹ năng ngôn ngữ và học tập cơ bản, bản sửa đổi
ADOS-2	Autism Diagnostic Observation Schedule, Second Edition	Chẩn đoán lịch trình quan sát chẩn đoán tự kỷ bản thứ 2
AOI	Area of Interest	Vùng quan tâm
BiLSTM	Bidirectional Long Short-Term Memory	Mạng bộ nhớ dài - ngắn hạn hai chiều
BKT	Bayesian Knowledge Tracing	Mô hình truy vết tri thức Bayes
CI	Confidence Interval	Khoảng tin cậy
CNN	Convolutional Neural Network	Mạng nơ-ron tích chập
CVAT	Computer Vision Annotation Tool	Công cụ chú thích thị giác máy tính
DAG	Directed Acyclic Graph	Đồ thị có hướng không chu trình
DBN	Dynamic Bayesian Network	Mạng Bayes động
DCM	Deterministic Centroid Minimization	Cực tiểu hóa tâm xác định
DIDUNAS	Digital Identification and Support of Under - Achieving Students	Hệ thống phát hiện và hỗ trợ kỹ thuật số cho học sinh gặp khó khăn trong học tập

DSM-5	Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition	Sổ tay chẩn đoán và thống kê các rối loạn tâm thần, phiên bản 5
fBm	Fractional Brownian Motion	Chuyển động Brownian phân số
FEM	Fixation Eye Movements	Vi chuyển động mắt
GA	Genetic Algorithm	Thuật toán di truyền
GRU	Gated Recurrent Unit	Mạng nơ-ron hồi quy với nút có cổng
I-DT	Identification by Dispersion-Threshold	Nhận dạng dựa trên ngưỡng phân tán
I-VT	Identification by Velocity-Threshold	Nhận dạng dựa trên ngưỡng vận tốc
IoU	Intersection over Union	Tỷ lệ diện tích giao cắt trên hợp
IR	Infrared light	Ánh sáng hồng ngoại
JSD	Jensen Shannon Divergence	Độ phân kỳ Jensen Shannon
KDE	Kernel Density Estimation	Ước lượng mật độ hạt nhân
LOOCV	Leave-One-Out Cross-Validation	Kiểm chứng chéo theo cấp độ đối tượng
LOSOCV	Leave-One-Subject-Out Cross-Validation	Kiểm chứng chéo luân phiên từng đối tượng
LSTM	Long Short-Term Memory	Mạng bộ nhớ dài - ngắn hạn
MLP	Multilayer Perceptron	Mạng nơ-ron đa lớp
OCR	Optical Character Recognition	Nhận dạng ký tự quang học
PCCR	Pupil Center Corneal Reflection	Tâm đồng tử và phản xạ giác mạc
PECS	Picture Exchange Communication System	Hệ thống giao tiếp trao đổi hình ảnh
PeyeMMV	Python implementation of eye movements metrics & visualizations	Thuật toán nhận diện điểm cố định

PSD	Power Spectral Density	Mật độ phổ công suất
PTDH		Phát triển điển hình
RFE	Recursive Feature Elimination	Chọn lọc đặc trưng đệ quy
RLPTK		Rối loạn phổ tự kỷ
RLTT		Rối loạn tính toán
RNN	Recurrent Neural Network	Mạng nơ-ron hồi quy
SAM	Segment Anything Model	Mô hình phân đoạn bất kì
SCED	Single Case Experimental Design	Thiết kế thực nghiệm trường hợp đơn
SGG	Stochastic Gaussian Generation	Sinh ngẫu nhiên Gaussian
SLD	Specific Learning Disorder	Rối loạn học tập cụ thể
SVM	Support Vector Machine	Máy học véc tơ hỗ trợ
TAD	Task Analysis Diagram	Đồ thị phân tích phụ thuộc nhiệm vụ
TEACCH	Treatment and Education of Autistic and Related Communication Handicapped Children	Phương pháp giáo dục và trị liệu cho trẻ tự kỷ và trẻ có khó khăn về giao tiếp
TSTR	Train on Synthetic, Test on Real	Huấn luyện tổng hợp, kiểm tra thực tế
TTFF	Time to First Fixation	Thời gian đến điểm cố định đầu tiên

DANH MỤC BẢNG

2.1	Đặc điểm nhân khẩu học và lâm sàng của trẻ tham gia	36
2.2	Bộ dữ liệu chuyển động mắt của trẻ rối loạn phổ tự kỉ	38
2.3	Đặc điểm nhân khẩu học và lâm sàng của trẻ tham gia	40
2.4	Thống kê bộ dữ liệu đa phương thức của trẻ rối loạn tính toán	44
3.1	Kết quả hiệu năng chẩn đoán của hệ thống đề xuất qua kiểm chứng chéo LOOCV	69
3.2	Đánh giá đối chứng hiệu năng phân loại giữa cấu hình cơ sở và cấu hình lai	74
3.3	Đối sánh hiệu năng với các nghiên cứu phân loại RLPTK tiên tiên	75
3.4	Giá trị cận trên, cận dưới và trung bình của ma trận chuyển đổi ánh nhìn theo nhóm	77
4.1	Bảng tổng hợp: Cơ sở ánh xạ từ tín hiệu thị giác sang hành động thích ứng của hệ thống	94
4.2	Các chỉ số, cơ chế suy luận logic và diễn giải hành vi trong hệ thống hỗ trợ can thiệp cho trẻ rối loạn phát triển	101
4.3	Chỉ số hành vi thị giác và phân tích kiểm định tích Tau-U (SCED 80 phiên)	104
4.4	Các thành phần tùy chỉnh hệ thống và mục tiêu hỗ trợ nhận thức	115
4.5	Kết quả đánh giá trước và sau can thiệp của 5 học sinh (thực nghiệm vòng lặp trong)	120
4.6	Phản hồi trong bài đánh giá trước can thiệp	123
4.7	Giá trị $P(L)$ của các kỹ năng tiên quyết	124

DANH MỤC HÌNH VẼ

1.1	Phương pháp trung tâm đồng tử-phản xạ giác mạc	10
1.2	Đường quét mắt (trái) và bản đồ nhiệt (phải) thể hiện sự phân bổ chú ý của người dùng trong quá trình giải toán	12
1.3	Quy trình tổng quan tài liệu có hệ thống	17
2.1	Mô hình 4 bước tổng quát xây dựng bộ dữ liệu đa phương thức	31
2.2	Nhiệm vụ đánh giá thiên hướng chú ý giữa yếu tố xã hội và phi xã hội	34
2.3	Nhiệm vụ đánh giá phản ứng chú ý thị giác trước kích thích bất ngờ	34
2.4	Video dữ liệu kích thích thị giác chứa 12 hình ảnh cho mỗi phiên	35
2.5	Nhiệm vụ không gian thị giác	39
2.6	Nhiệm vụ định vị trực số	39
2.7	Nhận biết số lượng nhanh	39
2.8	Phối hợp thị giác - thính giác	39
2.9	Hệ thống thu thập dữ liệu đa phương thức cho trẻ rối loạn tính toán	41
2.10	Quy trình bốn bước đồng bộ hóa dữ liệu đa nguồn	43
3.1	Sơ đồ luồng phương pháp trích xuất đặc trưng thị giác tự động sử dụng kiến trúc lai (SAM + Thuật toán luồng quang học) và thuật toán PeyeMMV.	52
3.2	Sơ đồ mô tả quy trình 3 bước của thuật toán PeyeMMV . . .	53
3.3	Quy trình phân đoạn và lan truyền vùng quan tâm (AOI) theo chuỗi thời gian	55
3.4	Quy trình xác thực khả năng nhận diện vùng quan tâm tự động	59
3.5	Hệ thống thu thập và trích xuất đặc trưng thị giác cho trẻ RLPTK	62

3.6	Hệ thống thu thập và trích xuất đặc trưng thị giác cho trẻ có rối loạn tính toán	62
3.7	Kết quả các đặc trưng thống kê của chuyển động mắt	64
3.8	Bản đồ nhiệt của trẻ RLPTK (cột phải) và trẻ phát triển điển hình (cột trái) khi quan sát các hình ảnh có yếu tố xã hội	65
3.9	Biểu đồ minh họa chuỗi thời gian của một đối tượng.	66
3.10	Sơ đồ luồng đánh giá thực nghiệm và tối ưu hóa mô hình phân loại RLPTK	68
3.11	Đường cong ROC của mô hình	70
3.12	Ma trận nhầm lẫn	70
3.13	Phân tích độ tương quan và phân bố của các dấu ấn sinh học hành vi	71
3.14	Đối sánh trực quan quỹ đạo đường quét trung bình nhóm.	72
3.15	Đường cong ROC đối chứng giữa cấu hình cơ sở và cấu hình lai đề xuất	72
3.16	Đối sánh ma trận nhầm lẫn giữa hai cấu hình phương pháp	73
3.17	Phân bố mật độ hạt nhân KDE của các đặc trưng phân tách tốt nhất	74
3.18	Bố cục chính trang bài tập.	76
3.19	Bản đồ nhiệt và đường quét biểu diễn hành vi quan sát trong bài toán ước lượng trên trục số.	78
3.20	Đường quét cho các chiến lược ước tính khác nhau của trẻ rối loạn tính toán	78
3.21	Bản đồ nhiệt và đường quét của trẻ rối loạn tính toán trong bài toán nghe và lựa chọn đáp án	79
3.22	Bản đồ nhiệt và đường quét trên các dữ liệu liên quan đến xử lý không gian thị giác	79
3.23	Kiến trúc hệ thống tăng cường dữ liệu di truyền và kiểm chứng TSTR.	82
3.24	Phân tích đối sánh độ phân tán không gian của các cụm điểm cố định tăng cường. Đỏ (DCM - Tính hoàn toàn, mất phương sai sinh lý), Xanh dương (SGG - Nhiều trắng không cấu trúc, dễ bị bộ lọc từ chối), và Xanh lá (GA - Phân bố sinh lý tự nhiên, duy trì phổ $1/f$).	84
3.25	Kiểm chứng thống kê về các chỉ số độ chân thực phổ. (a) GA đạt phân kỳ cực thấp. (b) GA tái tạo thành công động lực nhiều hồng $1/f$	85

3.26	Bản đồ nhiệt biểu diễn độ chính xác phân loại TSTR. Mapping cường độ màu sắc cô lập sự suy giảm nghiêm trọng của các phương pháp sinh dữ liệu không cấu trúc (SGG, DCM) trên các mô hình nhạy cảm ranh giới (SVM, MLP).	86
3.27	Phân bố hiệu năng đa chiều. Đa giác của GA thể hiện hình thái bao phủ tròn đều, ôm sát quỹ đạo của dữ liệu cơ sở trên toàn bộ 5 kiến trúc thuật toán.	87
4.1	Kiến trúc hệ thống can thiệp luồng kép phân tách luồng xử lý thị giác (cơ sở luật) và luồng nhận thức (mạng Bayes) . . .	92
4.2	Năm giai đoạn của chu trình can thiệp trẻ RLPTK tích hợp theo dõi chuyển động mắt	97
4.3	Biểu đồ phân tích thị giác chuỗi thời gian (SCED) của ba trẻ.	105
4.4	Ngành học hoặc lĩnh vực chuyên môn của người tham gia khảo sát	107
4.5	Số liệu thống kê mục tiêu nhận biết và xử lý số	109
4.6	Số liệu thống kê mục tiêu đếm số	109
4.7	Số liệu thống kê mục tiêu thực hiện phép tính	110
4.8	Số liệu thống kê mục tiêu hứng thú	110
4.9	Số liệu thống kê mục tiêu tập trung và chú ý	111
4.10	Mô hình kiến trúc tổng quan luồng dữ liệu thích ứng của hệ thống can thiệp ngôn ngữ.	111
4.11	Kết quả đạt được mục tiêu Đọc thành tiếng qua các buổi thực nghiệm – học sinh lớp 1	112
4.12	Kết quả đạt được mục tiêu Đọc thành tiếng qua các buổi thực nghiệm – học sinh lớp 2	112
4.13	Kết quả đạt được mục tiêu hứng thú qua các buổi thực nghiệm – học sinh lớp 2	112
4.14	Kết quả đạt được mục tiêu tập trung và chú ý qua các buổi thực nghiệm – học sinh lớp 1	113
4.15	Kết quả đạt được mục tiêu đọc hiểu qua các buổi thực nghiệm – học sinh lớp 1	113
4.16	Kết quả đạt được mục tiêu đọc hiểu qua các buổi thực nghiệm – học sinh lớp 2	114
4.17	Kết quả đạt được mục tiêu hứng thú qua các buổi thực nghiệm – học sinh lớp 1	114
4.18	Kết quả đạt được mục tiêu tập trung và chú ý qua các buổi thực nghiệm – học sinh lớp 2	115

4.19 Vòng lặp trong: Cơ chế hạ định dạng	118
4.20 Vòng lặp ngoài: Truy vết ngược trên TAD	119

MỞ ĐẦU

1 Đặt vấn đề

Công nghệ theo dõi chuyển động mắt (eye-tracking) là phương pháp xác định và ghi nhận những chuyển động của mắt khi con người quan sát từ vị trí này sang vị trí khác [1]. Dựa trên khả năng này, công nghệ theo dõi chuyển động mắt hỗ trợ phân tích mối liên hệ giữa các biểu hiện *hành vi thị giác*¹ và tiến trình nhận thức của não bộ [2]. Về mặt kĩ thuật, giải pháp này chuyển đổi từ phương thức tiếp xúc, trong đó yêu cầu cố định vị trí đầu của người tham gia, sang phương thức đo lường không xâm lấn, giảm thiểu các ràng buộc về tư thế. Sự chuyển dịch này được hỗ trợ bởi cơ chế kết hợp giữa ánh sáng hồng ngoại cùng các cảm biến quang học, cho phép hệ thống thu nhận dữ liệu điểm nhìn mà không cần tiếp xúc vật lý với người tham gia [3, 4]. Việc tinh gọn cấu trúc phần cứng từ các hệ thống cố định sang các thiết bị di động hoặc tích hợp đã góp phần nâng cao tính khả dụng, từ đó mở rộng khả năng triển khai công nghệ trong các môi trường tương tác thực tế.

Sự chuyển dịch sang các giải pháp công nghệ di động và khả năng tương tác không xâm lấn đã mở rộng phạm vi triển khai của công nghệ theo dõi chuyển động mắt trong nhiều lĩnh vực [6]. Trong lĩnh vực y tế, công nghệ này cung cấp nguồn dữ liệu sinh lý khách quan, hỗ trợ quá trình chẩn đoán sớm các rối loạn thần kinh [7, 8]. Trong lĩnh vực thương mại, công nghệ này đóng vai trò là công cụ phân tích hành vi người dùng qua dữ liệu trực quan, hỗ trợ tối ưu hóa phương thức trình bày thông tin của doanh nghiệp [9]. Trong lĩnh vực giáo dục, công nghệ này cung cấp cơ chế đo lường trạng thái nhận thức, phản ánh khách quan mức độ chú ý thị giác và năng lực xử

1. *Hành vi thị giác* là một quá trình thu nhận thông tin có định hướng, trong đó ý định của não bộ và hoạt động quan sát của mắt phối hợp chặt chẽ để chọn lọc những thông tin cần thiết và phù hợp từ môi trường xung quanh [5].

lý thông tin của trẻ [10, 11]. Kế thừa nền tảng đo lường đó, nguồn dữ liệu thị giác được ứng dụng trực tiếp nhằm hỗ trợ phát hiện sớm các dấu hiệu phát triển không điển hình và nhận diện rào cản học tập, trong đó trọng tâm hướng đến nhóm trẻ rối loạn phát triển [12–14].

Rối loạn phát triển thần kinh là nhóm tình trạng xuất hiện sớm từ thời thơ ấu, gây suy giảm hoạt động chức năng của cá nhân, xã hội, học tập hoặc nghề nghiệp [15, 16]. Nhóm này bao gồm ba dạng rối loạn tiêu biểu: rối loạn phổ tự kỉ, chứng khó đọc và rối loạn tính toán [9, 17, 18]. Mỗi dạng rối loạn đặc trưng bởi những rào cản nhận thức và hành vi riêng biệt. Cụ thể: nhóm trẻ rối loạn phổ tự kỉ gặp rào cản trong quá trình giao tiếp xã hội và có xu hướng thể hiện các hành vi rập khuôn [19]; nhóm trẻ rối loạn tính toán gặp rào cản trong quá trình tiếp cận khái niệm số lượng và xử lý quy luật logic [20, 21]; nhóm trẻ mắc chứng khó đọc gặp rào cản trong quá trình nhận diện từ vựng và giải mã ngôn ngữ [17, 22]. Những đặc điểm nhận thức này đặt ra yêu cầu cấp thiết đối với các phương pháp phát hiện và can thiệp sớm [23], trong đó công nghệ theo dõi chuyển động mắt đóng vai trò là một phương thức đo lường và hỗ trợ khả thi [24–26].

Dựa trên cơ sở đo lường này, dữ liệu hành vi thị giác thể hiện những đặc trưng riêng biệt tương ứng với từng dạng rối loạn. Cụ thể, nhóm trẻ rối loạn phổ tự kỉ thể hiện xu hướng ưu tiên quan sát các đối tượng hình khối và suy giảm mức độ tập trung vào khuôn mặt người [27–30]. Nhóm trẻ mắc chứng khó đọc bộc lộ xu hướng kéo dài thời gian dừng mắt và gia tăng tần suất nhìn ngược lại các nội dung đã đọc phía trước [31–33]. Nhóm trẻ rối loạn tính toán phản ánh sự suy giảm năng lực duy trì chú ý vào vị trí mục tiêu và kéo dài thời lượng nhận diện các ký hiệu số học [34, 35]. Việc trích xuất và ghi nhận các chỉ số hành vi thị giác này hỗ trợ quy trình sàng lọc lâm sàng [36, 37] và cung cấp căn cứ đo lường khách quan hỗ trợ quá trình thiết kế các lộ trình can thiệp phù hợp [38, 39].

Mặc dù đã thể hiện được tiềm năng hỗ trợ trong lĩnh vực giáo dục đặc biệt, việc ứng dụng công nghệ theo dõi chuyển động mắt để hỗ trợ nhóm trẻ rối loạn phát triển vẫn đối mặt với những rào cản kỹ thuật và phương pháp luận quan trọng. Dựa trên mục tiêu thiết lập phương pháp hỗ trợ trẻ rối loạn phát triển, luận án hệ thống hóa các khoảng trống khoa học hiện tại thành ba nhóm bài toán chính cần giải quyết:

- Bài toán 1: Xây dựng bộ dữ liệu hành vi thị giác đa phương thức
- Các nghiên cứu trước đây chủ yếu tập trung thu thập dữ liệu tọa độ

điểm nhìn đối với các kích thích thị giác dạng tĩnh hoặc trong môi trường thực nghiệm được kiểm soát nghiêm ngặt [40, 41]. Các nghiên cứu hiện nay ghi nhận sự thiếu hụt các bộ dữ liệu đa phương thức tích hợp đồng bộ tọa độ điểm nhìn, âm thanh và thao tác tương tác trong môi trường kích thích bằng chuỗi khung hình động có *hiệu lực sinh thái*² cao. Hơn nữa, do thiếu các thông số phản ánh *tải nhận thức*³, quá trình phân tích gặp khó khăn khi xác định trọng số tương quan giữa hành vi thị giác và năng lực tư duy [42, 43]. Sự thiếu hụt này làm thiếu hụt các đặc trưng hành vi cần thiết và giảm khả năng tổng quát hóa của các mô hình trí tuệ nhân tạo

- Bài toán 2: Trích chọn đặc trưng không gian và thời gian tự động
Nhằm phân tích vùng quan tâm (AOI), các nghiên cứu truyền thống thường yêu cầu chuyên gia tiến hành gán nhãn thủ công hoặc thiết lập các ranh giới không gian tĩnh [46]. Khi áp dụng vào chuỗi khung hình động nơi đối tượng liên tục thay đổi hình thái và vị trí, các phương pháp xác định nhãn đối tượng trên khung hình tĩnh phát sinh sai số đáng kể, trong khi phương thức gán nhãn thủ công không đáp ứng yêu cầu xử lý thời gian thực [47]. Bên cạnh đó, sự giới hạn về quy mô bộ dữ liệu huấn luyện cũng đặt ra yêu cầu cấp thiết về việc phát triển kỹ thuật tăng cường dữ liệu. Do đó, việc thiết lập một quy trình tự động hóa nhằm bám sát đối tượng động và tăng cường dữ liệu đóng vai trò là bài toán kỹ thuật trọng tâm được luận án đề xuất phương pháp giải quyết tại Chương 3.
- Bài toán 3: Mô hình hóa nhận thức và cá nhân hóa cơ chế can thiệp
Khi xây dựng các hệ thống hỗ trợ can thiệp cho trẻ rối loạn phát triển, xu hướng phổ biến hiện nay là ứng dụng các kiến trúc mạng học sâu nhằm thực hiện bài toán phân loại trực tiếp dựa trên dữ liệu chuyển động mắt. Tuy nhiên, đối với đặc thù của lĩnh vực y tế và giáo dục đặc biệt, các mô hình học sâu thường bộc lộ hạn chế nhất định do thiếu tính minh bạch trong quá trình đưa ra quyết định [48]. Do đó, khoảng trống nghiên cứu được xác định là sự khuyết thiếu một kiến trúc ra

2. *Tính hiệu lực sinh thái* đề cập đến mức độ mà các điều kiện, tác nhân kích thích và hành vi được quan sát trong môi trường thử nghiệm gần giống với các điều kiện, tác nhân kích thích và hành vi trong bối cảnh thực tế [44].

3. *Tải nhận thức* được hiểu là tổng mức độ tập trung và năng lượng trí não mà một người cần huy động để tiếp nhận, xử lý thông tin tại một thời điểm. Khái niệm này phản ánh giới hạn về khả năng gánh vác các nhiệm vụ tư duy của con người, trong đó mỗi hoạt động trí tuệ đều chiếm dụng một khoảng không gian nhất định trong khả năng ghi nhớ tạm thời. Khi các yêu cầu của tác vụ vượt quá ngưỡng năng lực này, việc xử lý thông tin sẽ trở nên kém hiệu quả [45].

quyết định lai, có khả năng kết hợp tính minh bạch của hệ cơ sở luật với năng lực suy diễn của các mô hình xác suất động. Việc ứng dụng cơ sở xác suất cập nhật liên tục theo thời gian thực đóng vai trò tiền đề để hệ thống tự động cá nhân hóa lộ trình can thiệp. Mục tiêu kỹ thuật này được luận án định hình và đề xuất phương pháp giải quyết tại Chương 4.

2 Mục tiêu nghiên cứu

Từ những thách thức nêu trên, luận án hướng tới nghiên cứu và phát triển giải pháp công nghệ, trong đó trọng tâm là công nghệ theo dõi chuyển động mắt nhằm hỗ trợ trẻ rối loạn phát triển. Mục tiêu tổng quát này được cụ thể hóa thành ba mục tiêu nghiên cứu thành phần, mỗi mục tiêu tập trung giải quyết một thách thức nêu trên:

- Mục tiêu 1: Xây dựng bộ dữ liệu đa phương thức đồng bộ theo thời gian thực (chuyển động mắt, tiếng nói, tương tác) phục vụ cho việc huấn luyện và kiểm chứng các mô hình học máy nhằm phân tích hành vi của trẻ rối loạn phát triển để hỗ trợ phân loại trẻ có nguy cơ rối loạn và xây dựng lộ trình can thiệp cá nhân hoá.
- Mục tiêu 2: Đề xuất các phương pháp xử lý và tăng cường dữ liệu chuyển động mắt nhằm cải thiện chất lượng thông tin đầu vào cho các mô hình phân tích. Đóng góp này bao gồm hai khía cạnh bổ trợ: (1) Đưa ra được phương pháp để phân đoạn dữ liệu trên khung hình động; (2) Đề xuất phương pháp giải quyết bài toán thiếu dữ liệu.
- Mục tiêu 3: Đề xuất phương pháp hỗ trợ can thiệp cá nhân hóa dựa trên dữ liệu hành vi thị giác được trích xuất từ dữ liệu theo dõi chuyển động mắt và phản hồi tương tác thu nhận qua hệ thống số hoá. Trọng tâm của mục tiêu này là xây dựng các quy tắc hỗ trợ ra quyết định điều chỉnh can thiệp hoặc điều phối nội dung học liệu và tốc độ tương tác của trẻ phù hợp với đặc điểm nhận thức của từng trẻ.

3 Đóng góp của luận án

Những đóng góp chính của luận án bao gồm:

- Xây dựng hai bộ dữ liệu đa phương thức của nhóm trẻ rối loạn phổ tự kỉ và nhóm trẻ rối loạn tính toán. Dữ liệu bao gồm các thành phần được thu thập đồng bộ theo thời gian thực: dữ liệu chuyển động mắt (tọa độ điểm nhìn, thời gian dừng mắt, video ghi màn hình, tỷ lệ hiển thị màn hình, mốc thời gian của từng khung hình), dữ liệu âm thanh (âm thanh hệ thống và âm thanh của người dùng trong suốt quá trình thu) và dữ liệu tương tác hệ thống (tọa độ chuột, phản hồi sự kiện, thời gian thao tác chuột). Các kích thích thị giác được thiết kế nhằm xác định đặc điểm nhận thức và phương thức xử lý thông tin của trẻ. Sau quá trình thu nhận, chuỗi dữ liệu được tiền xử lý và đồng bộ hóa theo thời gian, tạo cơ sở để huấn luyện các mô hình học máy phục vụ hỗ trợ sàng lọc trẻ có nguy cơ mắc các rối loạn phát triển, đồng thời hỗ trợ thiết kế các kế hoạch can thiệp giáo dục.
- Đề xuất phương pháp trích xuất đặc trưng thị giác tự động từ dữ liệu theo dõi chuyển động mắt, nhằm giải quyết rào cản kỹ thuật khi xử lý các kích thích dạng chuỗi khung hình động và giải quyết vấn đề mất cân bằng dữ liệu. Để thực hiện mục tiêu này, luận án tích hợp các mô hình phân đoạn đối tượng nhằm nhận diện và phân tách tự động các đối tượng mục tiêu trong từng khung hình, tạo cơ sở để thiết lập các vùng quan tâm biến đổi theo vị trí đối tượng. Dựa trên các vùng quan tâm đã được thiết lập tự động, dữ liệu chuyển động mắt được ánh xạ tương ứng vào từng đối tượng đã được định danh, giúp trích xuất các thông số phản ánh quá trình tiếp nhận thị giác và xử lý thông tin của trẻ. Những đặc trưng thị giác này cung cấp nguồn dữ liệu đầu vào cần thiết cho các mô hình học máy, phục vụ mục đích mô hình hóa hành vi nhận thức và xây dựng lộ trình hỗ trợ cá nhân hóa can thiệp trong giai đoạn tiếp theo. Bên cạnh đó, luận án đề xuất phương pháp tăng cường dữ liệu quỹ đạo ánh nhìn dựa trên thuật toán tối ưu tiến hóa (GA) và mô hình chuyển động Brownian phân số (fBm) nhằm khắc phục giới hạn về quy mô mẫu. Tập hợp các kỹ thuật này thực hiện chức năng chuyển đổi dữ liệu thô thành tập hợp các đặc trưng, phục vụ trực tiếp cho các bài toán phân loại ở giai đoạn sau.
- Đề xuất các giải pháp công nghệ, trọng tâm là công nghệ theo dõi chuyển động mắt, nhằm hỗ trợ can thiệp cá nhân hóa cho trẻ rối loạn phát triển. Theo hướng tiếp cận này, các đặc trưng hành vi thị giác được trích xuất và phân tích tạo cơ sở thực chứng để điều chỉnh các phương thức hỗ trợ, nhằm đáp ứng phù hợp với đặc điểm nhận thức và

nhu cầu tương tác riêng biệt của trẻ. Giải pháp đề xuất có khả năng tích hợp vào các nền tảng công nghệ giáo dục thực tế, góp phần điều chỉnh lộ trình can thiệp và hỗ trợ cải thiện năng lực cho trẻ rối loạn phát triển.

Các kết quả nghiên cứu của luận án đã được công bố và báo cáo tại các tạp chí, hội thảo và diễn đàn khoa học, bao gồm:

- 03 bài báo đăng trên các tạp chí quốc tế thuộc danh mục Web of Science (WoS) và Scopus, được xếp hạng Q1
- 04 bài báo đăng trong kỷ yếu hội thảo/hội nghị khoa học quốc tế
- 01 bài báo đăng trên tạp chí khoa học trong nước
- 01 báo cáo được trình bày tại diễn đàn khoa học cấp châu Á
- 01 đơn đăng ký sáng chế đã được chấp nhận

Ngoài ra, 03 bài báo gửi đăng trên các tạp chí quốc tế thuộc danh mục Web of Science và Scopus, được xếp hạng Q1, hiện đang trong quá trình xét duyệt.

4 Cấu trúc luận án

Nội dung tiếp theo của luận án được tổ chức như sau:

- Chương 1: Chương này trình bày các kiến thức nền tảng về công nghệ theo dõi chuyển động mắt và các hướng nghiên cứu ứng dụng công nghệ này trong việc phân tích đặc trưng thị giác của trẻ rối loạn phát triển. Nội dung chương tập trung tổng hợp và phân tích các công trình nghiên cứu hiện có nhằm cung cấp cơ sở lý luận và thực tiễn về những phương pháp tiếp cận đã được triển khai. Trên cơ sở đó, các khoảng trống nghiên cứu được xác định, làm tiền đề đề xuất hướng tiếp cận phù hợp cho luận án.
- Chương 2: Chương này tập trung xây dựng các bộ dữ liệu đa phương thức của trẻ rối loạn phát triển. Cụ thể, nội dung chương trình bày quy trình thiết kế và triển khai hệ thống thu thập dữ liệu. Kết quả xây dựng được hai bộ dữ liệu đa phương thức của nhóm trẻ rối loạn phổ tự kỷ và nhóm trẻ rối loạn tính toán cung cấp cơ sở định lượng cho các nghiên cứu phân tích đặc trưng hành vi thị giác, từ đó hỗ trợ quy trình sàng lọc trẻ có nguy cơ mắc các rối loạn phát triển tương ứng.

- Chương 3: Chương này tập trung phát triển các phương pháp xử lý và tăng cường dữ liệu chuyển động mắt. Nội dung chương mô tả quy trình xây dựng phương pháp lai tích hợp mô hình SAM và thuật toán PeyeMMV nhằm trích xuất tự động đặc trưng thị giác trên chuỗi khung hình động. Đồng thời, chương trình bày phương pháp tăng cường dữ liệu dựa trên thuật toán tối ưu tiến hóa nhằm khắc phục giới hạn về quy mô mẫu. Kết quả của quy trình này là một bộ dữ liệu có cấu trúc và đồng bộ.
- Chương 4: Chương này tập trung khai thác các đặc trưng thị giác đã trích xuất làm đầu vào cho các mô hình tính toán, nhằm mục đích hỗ trợ nhận diện rào cản học tập và tối ưu hóa lộ trình can thiệp cá nhân hóa cho trẻ rối loạn phát triển. Nhằm nâng cao tính thực tiễn của nghiên cứu, luận án đề xuất các kiến trúc hệ thống được thiết kế dựa trên năng lực tương tác và quy trình xử lý thông tin của từng nhóm trẻ rối loạn phát triển. Các kiến trúc này được cụ thể hóa qua hai giải pháp: một hệ thống nền tảng web hỗ trợ học toán cho trẻ rối loạn tính toán và một ứng dụng di động hỗ trợ đọc, xử lý văn bản cho trẻ mắc chứng khó đọc.
- Phần kết luận: Phần này tổng hợp kết quả của nghiên cứu đã đạt được trong quá trình thực hiện luận án đồng thời bàn luận về những giới hạn của phương pháp đề xuất và định hướng nghiên cứu.

Chương 1

TỔNG QUAN VỀ CÔNG NGHỆ THEO DÕI CHUYỂN ĐỘNG MẮT HỖ TRỢ TRẺ RỐI LOẠN PHÁT TRIỂN

Chương 1 trình bày những kiến thức cơ sở phục vụ cho quá trình nghiên cứu và đề xuất giải pháp ứng dụng công nghệ theo dõi chuyển động mắt trong hỗ trợ trẻ rối loạn phát triển. Nội dung chương bao gồm tổng quan về công nghệ theo dõi chuyển động mắt và các đặc trưng chuyển động mắt ở từng nhóm rối loạn phát triển. Bên cạnh đó, luận án tổng hợp và hệ thống hóa các công trình đã công bố liên quan đến việc ứng dụng công nghệ này cho các nhóm trẻ rối loạn phát triển khác nhau, từ đó rút ra những kết quả đạt được, đồng thời chỉ ra các hạn chế, thách thức và định hướng nghiên cứu.

1.1 Công nghệ theo dõi chuyển động mắt

Công nghệ theo dõi chuyển động mắt là phương pháp ghi nhận và phân tích chuyển động của mắt khi con người chuyển hướng nhìn giữa các vị trí khác nhau [1]. Thông qua đó, công nghệ này giúp làm rõ mối liên hệ giữa hoạt động của chuyển động mắt và hệ thần kinh, từ đó liên kết hành vi nhìn với quá trình nhận thức và cung cấp thông tin về hoạt động bên trong của não bộ [2]. Các nghiên cứu về chuyển động mắt bắt đầu hình thành từ

thế kỷ XIX. Tuy nhiên, những phương pháp sơ khai này thường mang tính xâm lấn do việc gắn trực tiếp thấu kính lên giác mạc để ghi nhận dữ liệu, dẫn đến những ảnh hưởng đến sự thoải mái cho người tham gia [49, 50]. Dù còn hạn chế, các phương pháp này vẫn tạo cơ sở trong việc ghi nhận chuyển động mắt, từ đó đặt nền tảng cho các nghiên cứu tiếp theo nhằm phát triển những kỹ thuật theo dõi không xâm lấn. Ngày nay, chuyển động mắt có thể được ghi nhận thông qua các thiết bị theo dõi ánh nhìn không xâm lấn. Về cơ bản, phần lớn các thiết bị này hoạt động bằng cách chiếu một nguồn sáng lên bề mặt mắt và ghi lại tín hiệu phản xạ thu được [3]. Nguồn sáng thường là điốt hồng ngoại (infrared light - IR light), kết hợp với camera hướng về phía mắt của người tham gia [4, 51]. Ánh sáng hồng ngoại này không thể nhìn thấy bằng mắt thường, nhưng tạo ra các điểm phản xạ trên bề mặt mắt và được camera ghi lại.

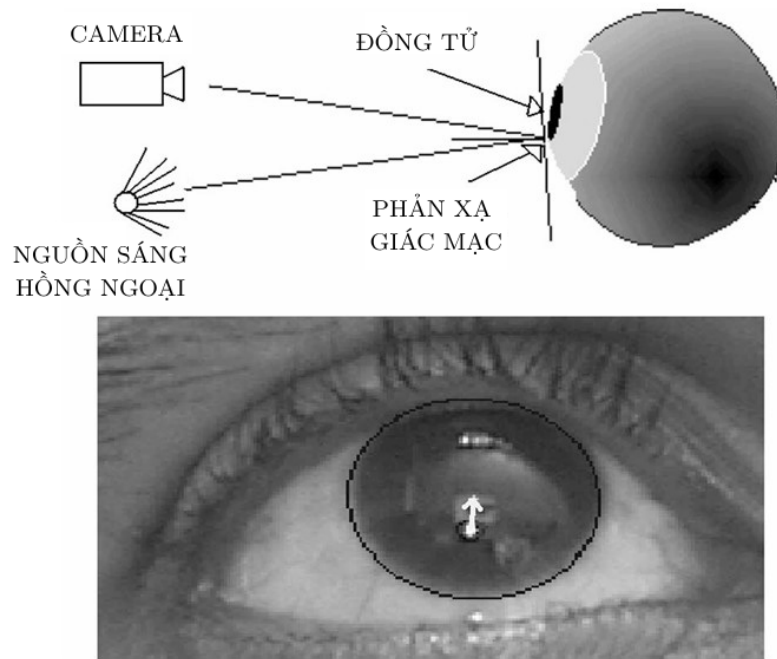
Trong số các kỹ thuật được sử dụng để theo dõi chuyển động mắt, phương pháp tâm đồng tử và phản xạ giác mạc (Pupil center corneal reflection - PCCR) là kỹ thuật phổ biến và được tích hợp rộng rãi trong các thiết bị theo dõi chuyển động mắt hiện đại [3], được minh họa trong Hình 1.1. Tại mỗi bước thời gian t , hệ thống kết xuất một véc-tơ trạng thái sinh trắc học [52]:

$$S_t = [x_t, y_t, d_t, \tau_t]^T \quad (1.1)$$

Trong đó:

- (x_t, y_t) là tọa độ không gian 2D của điểm nhìn trên mặt phẳng giao diện màn hình.
- d_t là đường kính đồng tử (pupil size), phản ánh gián tiếp mức độ tải nhận thức.
- τ_t là nhãn thời gian (timestamp) tương ứng với chu kỳ lấy mẫu của thiết bị.

Chuỗi dữ liệu tổng hợp $G = \{S_1, S_2, \dots, S_T\}$ là dữ liệu thô (raw) và thường chứa nhiều sai số do các chuyển động không chủ đích. Việc thiết kế các thuật toán nhằm chuyển đổi không gian véc-tơ thô này thành các đặc trưng có ý nghĩa chính là tiền đề kỹ thuật quan trọng cho các mô hình phân tích và học máy ở các giai đoạn sau.



Hình 1.1: Phương pháp trung tâm đồng tử-phản xạ giác mạc

1.1.1 Các đặc trưng của dữ liệu chuyển động mắt

Trong các bài toán liên quan đến chuyển động mắt, việc nghiên cứu và lựa chọn các đặc trưng phù hợp đóng vai trò quan trọng [51]. Phần lớn các thiết bị theo dõi chuyển động mắt hiện nay thu nhận dữ liệu dưới dạng chuỗi các điểm nhìn, được biểu diễn bằng tọa độ (x, y) kèm theo mốc thời gian tương ứng [52]. Tuy nhiên, do mật độ điểm nhìn ghi nhận được thường lớn, việc sử dụng trực tiếp nguồn dữ liệu này gây ra những rào cản nhất định trong quá trình phân tích và phát hiện các đặc điểm nhận thức quan trọng [53]. Vì vậy, tùy thuộc vào mục tiêu nghiên cứu, dữ liệu điểm nhìn thường được xử lý để trích xuất các đặc trưng chuyển động mắt có ý nghĩa hành vi, như điểm cố định, đường trượt và hồi quy, giúp quá trình quan sát và đánh giá trở nên hiệu quả hơn [54].

Điểm cố định là đặc trưng biểu thị khoảng thời gian mà mắt duy trì trạng thái nhìn cố định vào một mục tiêu thị giác cụ thể, nhằm phục vụ cho quá trình nhận thức và tiếp nhận thông tin thị giác [32, 55]. Việc xác định các điểm cố định cho phép đơn giản hóa dữ liệu chuyển động mắt bằng cách gom các điểm nhìn rời rạc thành những đơn vị biểu diễn có ý nghĩa hơn [54]. Trong khi đó, đường trượt là dạng chuyển động nhanh của mắt khi di chuyển từ một điểm cố định sang điểm cố định tiếp theo [2]. Theo

Fischer, con người thường thực hiện trung bình từ 3 đến 5 đường trượt mỗi giây [56]. Điểm cố định và đường trượt là hai cơ chế giúp con người tương tác và nhận thức thế giới: trong khi điểm cố định là giai đoạn diễn ra quá trình xử lý và tiếp nhận thông tin, thì đường trượt đóng vai trò dẫn hướng ánh nhìn đến các vị trí chứa thông tin tiếp theo [32, 55]. Vì vậy, hai đặc trưng này là những yếu tố được sử dụng phổ biến trong các nghiên cứu ứng dụng công nghệ theo dõi chuyển động mắt [52].

Bên cạnh điểm cố định và đường trượt, hồi quy cũng là một dạng chuyển động mắt quan trọng. Hồi quy được hiểu là chuyển động đường trượt theo hướng ngược lại với tiến trình quan sát thông thường, tức là mắt quay trở lại vị trí đã quan sát trước đó [32, 50]. Hồi quy thường được áp dụng trong các bài toán nghiên cứu về hoạt động đọc và quá trình xử lý thông tin [55]. Hồi quy khi đọc có thể xuất phát từ nhiều nguyên nhân. Trước hết, trong quá trình đọc, mắt nhìn có thể bỏ sót những thông tin quan trọng hoặc rời khỏi điểm cố định quá sớm, khiến việc xử lý thông tin chưa hoàn tất và gây ra sự không chắc chắn trong việc tiếp nhận nội dung [32]. Bên cạnh đó, hồi quy còn gắn liền với năng lực trí nhớ công việc, đóng vai trò hỗ trợ người đọc hồi tưởng lại những thông tin đã đọc trước đó, đặc biệt trong trường hợp các văn bản dài và đòi hỏi khả năng xử lý thông tin cao [32, 55].

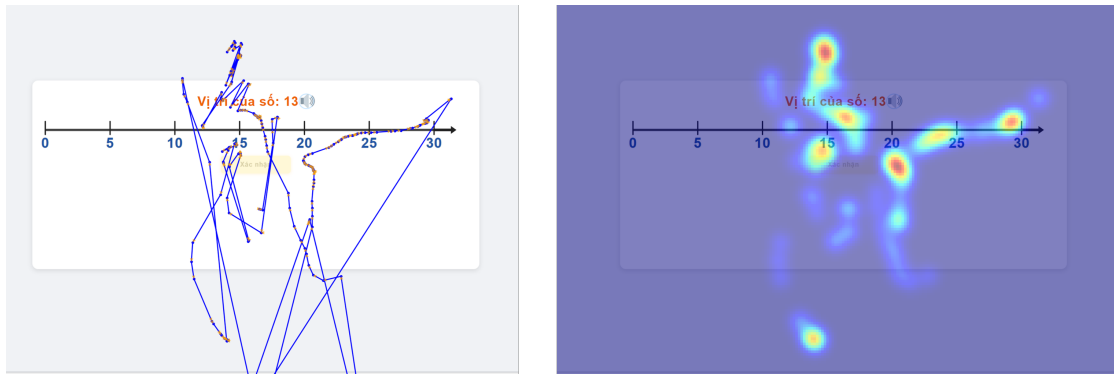
1.1.2 Các phương pháp trực quan hóa dữ liệu chuyển động mắt

Các đặc trưng của dữ liệu chuyển động mắt được trình bày tại mục 1.1.1 chủ yếu mang tính chất thống kê. Để hỗ trợ việc phân tích và diễn giải các đặc trưng này một cách trực quan, các phương pháp như bản đồ nhiệt và đường quét thường được sử dụng [52, 57].

Bản đồ nhiệt (heat map) là một kỹ thuật dùng để biểu diễn sự phân bố không gian của chuyển động mắt trên toàn bộ kích thích thị giác [52]. Vì vậy, bản đồ nhiệt thường tập trung vào mật độ phân bố của các điểm cố định mà không xét đến thứ tự xuất hiện của chúng [57]. Khi trực quan hóa, một lớp mặt nạ màu được tích hợp lên hình ảnh kích thích mà người tham gia quan sát, trong đó màu sắc thể hiện mật độ điểm cố định tại từng vị trí. Các vùng có màu nóng biểu thị nơi có nhiều điểm cố định hơn, qua đó phản ánh mức độ tập trung chú ý cao của người quan sát tại những vùng đó [52].

Trong khi đó, đường quét là biểu diễn trực quan của chuỗi các điểm cố định và đường trượt, phản ánh quỹ đạo chuyển động của mắt trong quá trình thực hiện một nhiệm vụ thị giác [52, 58]. Thông thường, đường quét được thể hiện dưới dạng các nút (tương ứng với các điểm cố định) và các cạnh nối giữa chúng (biểu thị các đường trượt), trong đó kích thước của mỗi nút tỷ lệ thuận với thời lượng của điểm cố định tương ứng [58]. Do đặc tính quan sát tự nhiên của con người, mắt thường xuyên quay lại các vị trí đã nhìn trước đó, khiến cho các điểm cố định và đường trượt có xu hướng chồng chéo lên nhau. Vì vậy, đường quét thường có cấu trúc phức tạp, đòi hỏi phải được phân tích cẩn thận để rút ra các thông tin có ý nghĩa về hành vi thị giác của con người [57].

Hình 1.2 minh họa một ví dụ về đường quét và bản đồ nhiệt của trẻ trong quá trình giải toán.



Hình 1.2: Đường quét mắt (trái) và bản đồ nhiệt (phải) thể hiện sự phân bố chú ý của người dùng trong quá trình giải toán

1.1.3 Phương pháp xác định vùng quan tâm trong phân tích dữ liệu chuyển động mắt

Việc phân tích dữ liệu thị giác trên toàn bộ khung hình thường tiêu tốn tài nguyên và khó làm nổi bật các đặc trưng riêng biệt [46]. Do đó, các mô hình theo dõi chuyển động mắt thường áp dụng phương pháp gắn dữ liệu chuyển động mắt với một đối tượng mục tiêu [47]. Để thống nhất trong các phân tích, luận án xác lập hai khái niệm nền tảng như sau:

- Cố định thị giác: Là trạng thái mắt dừng lại và tập trung nhìn vào một vùng không gian cố định trên khung hình trong một khoảng thời gian nhất định (thường $\geq 100ms$), biểu thị quá trình não bộ đang xử lý thông tin tại vị trí đó [32].

- Vùng quan tâm (Area of Interest - AOI): Là vùng không gian hình học được xác định trên khung hình hiển thị kích thích, bao quanh các đối tượng kích thích (từ ngữ, câu văn, hình ảnh) nhằm đo lường mức độ phân bổ chú ý của trẻ. Khái niệm AOI quan trọng trong các ứng dụng hỗ trợ trẻ rối loạn phát triển, khi việc định hướng chú ý theo kích thích có liên hệ chặt chẽ với khả năng nhận thức [46, 59].

Khái niệm này được sử dụng trong các nghiên cứu hỗ trợ trẻ rối loạn phát triển, khi việc định hướng chú ý theo kích thích có liên hệ chặt chẽ với khả năng nhận thức của trẻ [47]. Tùy theo từng nhóm rối loạn phát triển, cách xác định vùng quan tâm cũng có sự khác biệt. Đối với trẻ rối loạn phổ tự kỉ (RLPTK), vùng quan tâm thường tập trung vào các vùng chứa hình ảnh đồ vật, con vật hoặc các kích thích thị giác nổi bật [60, 61]. Trong khi đó, đối với trẻ mắc chứng khó đọc và rối loạn tính toán (RLTT), vùng quan tâm thường được xác định tại các vùng chứa văn bản, ví dụ như toàn bộ đoạn văn (thường được mở rộng thêm khoảng 0,5 cm xung quanh), các câu riêng lẻ, hoặc các đơn vị nhỏ hơn như từ, chữ số hay phép tính [62].

1.2 Đặc trưng hành vi thị giác của trẻ rối loạn phát triển

Theo Sổ tay chẩn đoán và thống kê các loại rối loạn tâm thần tái bản lần 5 (Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition - DSM-5) [63], rối loạn phát triển thần kinh là nhóm tình trạng xuất hiện sớm từ thời thơ ấu, gây suy giảm chức năng cá nhân, xã hội, học tập hoặc nghề nghiệp [9, 64]. Tùy thuộc vào từng phổ rối loạn cụ thể, trẻ sẽ thể hiện các đặc trưng lâm sàng khác nhau. Đối với rối loạn phổ tự kỉ, khiếm khuyết trong giao tiếp xã hội và các hành vi định hình, lặp lại là những triệu chứng cốt lõi [65, 66]. Trong khi đó, đối với nhóm rối loạn học tập cụ thể (Specific Learning Disorder - SLD), trẻ mắc chứng khó đọc gặp khó khăn trong việc nhận diện và giải mã âm vị, còn trẻ rối loạn tính toán đối diện với rào cản trong xử lý không gian và ước lượng số lượng [9].

1.2.1 Rối loạn phổ tự kỉ

Rối loạn phổ tự kỉ được đặc trưng bởi những thiếu hụt dai dẳng trong giao tiếp và tương tác xã hội trên nhiều bối cảnh, kèm theo các hành vi, sở

thích hoặc hoạt động hạn chế và lặp lại, xuất hiện từ giai đoạn phát triển sớm [63]. Sự giao thoa giữa các đặc trưng này tạo ra những rào cản kép, gây ra những thách thức trong tiến trình học tập cũng như sinh hoạt hằng ngày của người mắc rối loạn này. Trong môi trường học tập, hệ thống rào cản này làm hạn chế khả năng tiếp nhận và xử lý thông tin; trong khi đó, tại môi trường sinh hoạt, chúng gây trở ngại cho việc giao tiếp và hòa nhập cộng đồng. Với tỷ lệ chẩn đoán đang có xu hướng gia tăng, việc phát triển các phương pháp nghiên cứu và giải pháp can thiệp riêng biệt trở thành một yêu cầu thực tiễn quan trọng [67].

Đối với nhóm trẻ RLPTK, sự khác biệt trong cách phân bổ và di chuyển ánh nhìn được thể hiện qua hành vi thực tế mà dữ liệu chuyển động mắt đo được. Đối với các hành vi quan sát được, trẻ thường tránh giao tiếp bằng mắt và chú ý nhiều đến các kích thích mạnh [68–70]. Trong khi đó, các đặc trưng chuyển động mắt cho thấy trẻ ít tập trung vào các yếu tố xã hội như khuôn mặt, mà thường nhìn lâu vào các vật thể hoặc chuyển động lặp lại [27, 30, 71]. Ngoài ra, trẻ có xu hướng duy trì ánh nhìn cố định kéo dài và phản xạ đồng tử chậm, gây khó khăn trong việc chuyển hướng chú ý [61, 72, 73].

1.2.2 Rối loạn tính toán

Theo DSM-5, thuật ngữ “bị khiếm khuyết về toán học” là một dạng rối loạn học tập, đặc trưng bởi những khó khăn dai dẳng trong các kỹ năng toán học như cảm giác số, ghi nhớ các sự kiện số học, thực hiện phép tính một cách chính xác hoặc trôi chảy, cũng như khả năng lý luận toán học [9]. Thông thường, bị khiếm khuyết về toán học được gọi bằng một thuật ngữ thay thế quen thuộc hơn là rối loạn tính toán [63]. Khác với các khó khăn học tập, sự thiếu hụt mà trẻ RLTT gặp phải chủ yếu liên quan đến lĩnh vực toán học và khó được khắc phục bằng các phương pháp giảng dạy thông thường trên lớp [74]. Trẻ thường gặp khó khăn trong việc nhận biết nhanh số lượng mà không cần đếm tuần tự từng đối tượng, khó khăn trong việc so sánh số lượng lớn - bé, hạn chế khả năng hình dung vị trí của các con số trên trục số trong não bộ [21]. Hiện nay, tỷ lệ mắc RLTT ước tính dao động từ 3% đến 6% dân số trong độ tuổi đi học [75]. Mặc dù có mức độ phổ biến tương đối cao, RLTT hầu như vẫn chưa nhận được sự quan tâm và nghiên cứu đầy đủ như các rối loạn phát triển khác [74].

Những khó khăn mà trẻ RLTT gặp phải được thể hiện qua chuyển động mắt khi nào bộ xử lý các con số, ký hiệu hay phép toán [34, 76]. Các nghiên cứu gần đây cho thấy, trẻ em mắc rối loạn này có kiểu chuyển động mắt khác biệt so với những trẻ phát triển điển hình (PTDH) cùng độ tuổi [34, 77]. Cụ thể, trẻ RLTT thường gặp tình trạng mã hóa số chậm do khả năng xử lý thông tin thị giác kém hiệu quả, dễ bị phân tán chú ý và có chuyển động mắt thiếu ổn định [34, 76]. Trong các nhiệm vụ nhận biết nhanh số lượng vật thể mà không cần đếm, thay vì sử dụng chiến lược nhận biết tức thì, trẻ thường phải đếm lần lượt từng đối tượng ít nhất một lần [34]. Điều này làm tăng thời gian phản ứng, đồng thời số lượng điểm cố định và thời lượng điểm cố định cũng gia tăng [10, 34]. Ngoài ra, trong các bài toán liên quan đến trục số, trẻ RLTT có thể gặp khó khăn khi lựa chọn cách ước lượng phù hợp [35]. Trẻ thường sử dụng những điểm tham chiếu nằm xa giá trị cần xác định do ước lượng ban đầu đã lệch khỏi mục tiêu [35]. Điều này thể hiện qua việc các điểm cố định mắt phân bố rải rác hơn trên trục số so với trẻ PTDH cùng độ tuổi [35, 77].

1.2.3 Chứng khó đọc

Theo DSM-5, thuật ngữ “bị suy giảm khả năng đọc” là một loại rối loạn học tập đặc hiệu, đặc trưng bởi những khó khăn dai dẳng trong các kỹ năng đọc quan trọng như độ chính xác khi đọc từ, tốc độ đọc hoặc mức độ trôi chảy, cũng như khả năng đọc hiểu. Thông thường, tình trạng này thường được gọi bằng thuật ngữ thay thế quen thuộc hơn là chứng khó đọc, dùng để chỉ một mô hình khó khăn trong học tập liên quan đến việc nhận dạng từ không chính xác hoặc thiếu trôi chảy, khả năng giải mã kém và hạn chế trong đánh vần [63]. Trẻ mắc chứng khó đọc thường gặp trở ngại trong việc ánh xạ giữa chữ cái và âm thanh, dẫn đến quá trình giải mã từ vựng diễn ra với tốc độ thấp và đòi hỏi tải nhận thức lớn. Những khó khăn này thường xuất phát từ sự thiếu hụt trong khả năng nhận diện các âm vị cấu thành ngôn ngữ [17, 78]. Hệ quả là trẻ phải đối mặt với những rào cản đáng kể trong quá trình đọc, thường biểu hiện qua các lỗi như đảo thứ tự chữ cái, bỏ sót từ hoặc khó khăn duy trì độ trôi chảy của văn bản [26, 36]. Những yếu tố này có thể gây ra ảnh hưởng đến kết quả học tập cũng như chất lượng cuộc sống hằng ngày của trẻ. Hiện nay, chứng khó đọc là vấn đề phổ biến trên toàn cầu với tỷ lệ mắc dao động từ 5% đến 10% ở độ tuổi đi học [79].

Trong các nghiên cứu về chuyển động mắt, trẻ mắc chứng khó đọc thể hiện những đặc trưng khác biệt so với trẻ PTĐH. Cụ thể, nhóm trẻ này thường có số lượng điểm cố định nhiều hơn và thời lượng điểm cố định kéo dài hơn, phản ánh nhu cầu về thời gian để xử lý thông tin, đồng thời cho thấy khả năng phối hợp vận động mắt kém hiệu quả [32, 36, 62, 80]. Ngoài ra, trẻ mắc chứng khó đọc thường có tần suất đường trượt cao hơn nhưng biên độ mỗi đường trượt lại ngắn hơn; điều này phản ánh hạn chế trong việc tiếp nhận thông tin thị giác cũng như sự giảm sút về mức độ tự tin khi đọc [32]. Bên cạnh đó, tần suất hồi quy cũng cao hơn, cho thấy năng lực đọc hiểu và trí nhớ công việc của trẻ gặp nhiều hạn chế so với nhóm trẻ PTĐH.

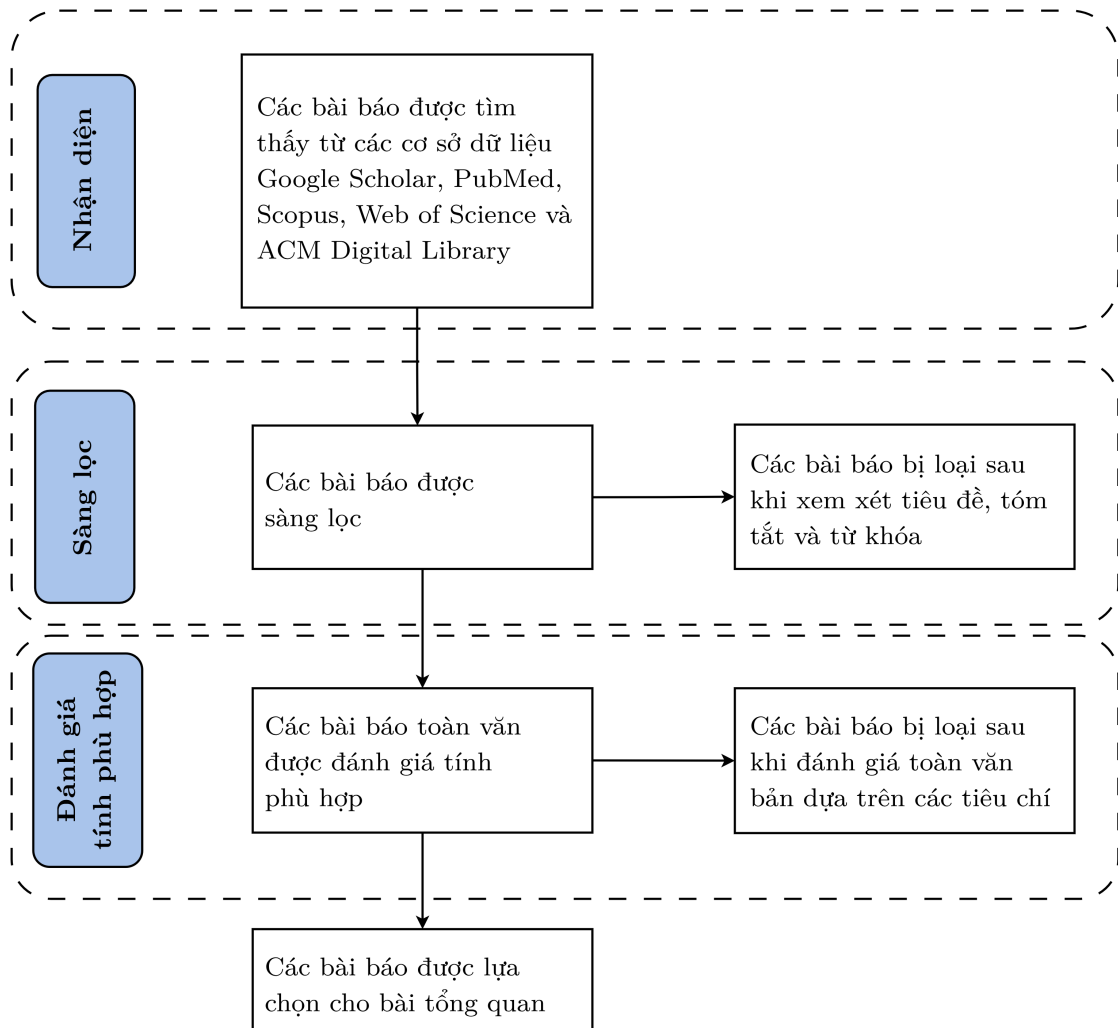
1.3 Hệ thống hóa giải pháp theo dõi chuyển động mắt hỗ trợ trẻ rối loạn phát triển

Trong những năm gần đây, công nghệ theo dõi chuyển động mắt là một hướng tiếp cận quan trọng trong nghiên cứu về hành vi thị giác và quá trình nhận thức của con người. Nhờ khả năng thu nhận dữ liệu ánh nhìn chi tiết và không xâm lấn, công nghệ này, khi kết hợp với các kỹ thuật thị giác máy tính và học máy, đã cung cấp nền tảng dữ liệu phục vụ cho việc phân tích định lượng các đặc trưng nhận thức [52]. Đặc biệt, nhờ những ưu điểm này, công nghệ theo dõi chuyển động mắt ngày càng được ứng dụng rộng rãi trong nghiên cứu và hỗ trợ trẻ rối loạn phát triển. Cụ thể, dữ liệu theo dõi chuyển động mắt được sử dụng để phân tích đặc trưng hành vi thị giác, xây dựng các mô hình phát hiện sớm, cũng như đánh giá hiệu quả can thiệp [13]. Tuy nhiên, sự gia tăng nhanh chóng về số lượng công trình nghiên cứu trong lĩnh vực này cũng đặt ra yêu cầu cần có các nghiên cứu tổng hợp nhằm hệ thống hoá lý thuyết, xác định các xu hướng phát triển, cũng như định hướng lựa chọn và phát triển các phương pháp phù hợp cho từng bài toán ứng dụng.

Từ thực tế đó, việc tổng quan và hệ thống hóa các công trình ứng dụng công nghệ theo dõi chuyển động mắt là cần thiết [71]. Vì vậy, luận án tập trung tổng hợp và phân tích các phương pháp tiếp cận dựa trên công nghệ này trong hỗ trợ trẻ rối loạn phát triển thông qua các tổng quan nghiên cứu có hệ thống, hướng tới các nhóm đối tượng bao gồm: Rối loạn phổ tự kỉ, rối loạn tính toán và chứng khó đọc.

1.3.1 Phương pháp luận

Phương pháp tổng quan tài liệu có hệ thống được sử dụng nhằm tổng hợp và phân tích các công trình nghiên cứu liên quan đến ứng dụng công nghệ theo dõi chuyển động mắt trong hỗ trợ trẻ rối loạn phát triển. Phương pháp này tuân theo quy trình do Kitchenham và Charters đề xuất [81], nhằm hỗ trợ việc lựa chọn tài liệu được khách quan, minh bạch và giảm thiểu sai lệch trong quá trình tổng hợp. Hình 1.3 minh họa ba giai đoạn chính của quy trình, bao gồm: xác định, sàng lọc và đánh giá tính phù hợp.



Hình 1.3: Quy trình tổng quan tài liệu có hệ thống

Cụ thể, trong giai đoạn xác định, các nghiên cứu được thu thập từ nhiều cơ sở dữ liệu khoa học như PubMed, Google Scholar, Scopus, Web of Science và ACM Digital Library. Các truy vấn tìm kiếm được xây dựng dựa trên các từ khóa liên quan đến công nghệ theo dõi chuyển động mắt và các rối loạn phát triển nói chung, bao gồm “*eye tracking*”, “*gaze behavior*”, kết hợp

với các từ khóa như “*autism*”, “*dyslexia*”, “*dyscalculia*” và các thuật ngữ liên quan đến “*diagnosis*”, “*intervention*”, “*analysis*”. Tiếp theo, ở giai đoạn sàng lọc, các nghiên cứu được xem xét sơ bộ dựa trên tiêu đề, tóm tắt và từ khóa. Trong quy trình này, sàng lọc được định nghĩa là quá trình loại bỏ các công trình không liên quan dựa trên các tiêu chí ban đầu, bao gồm: (i) không sử dụng dữ liệu theo dõi chuyển động mắt, (ii) không tập trung vào nhóm rối loạn phát triển hoặc (iii) không phải là bài báo khoa học (ví dụ, báo cáo kỹ thuật không phản biện hoặc bài viết phổ thông). Những nghiên cứu không đáp ứng các tiêu chí này sẽ bị loại bỏ nhằm duy trì tính nhất quán của bộ dữ liệu. Cuối cùng, trong giai đoạn đánh giá tính phù hợp, các nghiên cứu còn lại được đọc toàn văn để đánh giá theo các tiêu chí lựa chọn. Các tiêu chí này bao gồm: (i) mô tả rõ ràng phương pháp thu thập và phân tích dữ liệu theo dõi chuyển động mắt, (ii) có kết quả thực nghiệm và (iii) liên quan trực tiếp đến một trong ba mục tiêu: phân tích đặc trưng, hỗ trợ chẩn đoán hoặc hỗ trợ can thiệp. Các nghiên cứu trùng lặp hoặc không cung cấp đủ thông tin sẽ bị loại bỏ.

Trong khuôn khổ luận án, sau khi thực hiện quy trình tổng quan tài liệu có hệ thống, tổng cộng 170 nghiên cứu liên quan đến trẻ RLPTK và 169 nghiên cứu về RLTT đã được lựa chọn và phân tích. Bên cạnh đó, một tổng quan hệ thống trước đây về chứng khó đọc [26] cũng được đề cập nhằm mở rộng phạm vi phân tích. Dựa trên mục tiêu nghiên cứu, các công trình được phân loại thành ba hướng tiếp cận chính:

- Phân tích đặc trưng thị giác: Phân tích dữ liệu chuyển động mắt nhằm mô tả, làm rõ các đặc điểm hành vi thị giác cùng quá trình nhận thức của trẻ, tập trung vào các chỉ số như điểm cố định, đường trượt, hồi quy cũng như phân bố vùng quan tâm.
- Hỗ trợ chẩn đoán: Sử dụng dữ liệu chuyển động mắt kết hợp với các phương pháp học máy hoặc học sâu để phân loại và phát hiện sớm các rối loạn phát triển.
- Hỗ trợ can thiệp và đánh giá hiệu quả can thiệp: Ứng dụng công nghệ theo dõi chuyển động mắt trong thiết kế và đánh giá các chương trình can thiệp, cũng như sử dụng chuyển động mắt để đo lường hiệu quả can thiệp thông qua các thay đổi trong hành vi thị giác.

1.3.2 Công nghệ theo dõi chuyển động mắt trong phân tích đặc trưng thị giác

Hướng nghiên cứu ứng dụng công nghệ theo dõi chuyển động mắt trong phân tích đặc trưng hành vi thị giác của trẻ rối loạn phát triển là một hướng tiếp cận quan trọng, chiếm ưu thế lớn trong tổng số các nghiên cứu liên quan. Tùy theo từng nhóm rối loạn, dữ liệu chuyển động mắt có thể phản ánh những khía cạnh khác nhau trong hành vi và quá trình nhận thức của trẻ [24, 38].

Cụ thể, ở trẻ RLPTK, các đặc trưng này cho thấy sự thiếu hụt trong định hướng xã hội và xu hướng hành vi lặp lại, qua đó cung cấp các chỉ số định lượng khách quan hỗ trợ sàng lọc và đánh giá. Ví dụ, dữ liệu theo dõi chuyển động mắt cho thấy sự thiếu hụt trong định hướng xã hội của trẻ RLPTK, thể hiện qua việc giảm chú ý vào vùng mắt và tăng chú ý vào miệng hoặc các chi tiết không liên quan đến con người trong bối cảnh [27, 30]. Ngoài ra, trẻ còn thể hiện xu hướng tập trung kéo dài vào các mẫu hình học hoặc cấu trúc lặp lại [28].

Đối với trẻ có rối loạn tính toán, công nghệ theo dõi chuyển động mắt được sử dụng để làm rõ những thiếu hụt trong xử lý số học và sự khác biệt so với nhóm PTDH [82]. Trẻ thường sử dụng chiến lược đếm tuần tự ngay cả với các tập hợp nhỏ, thể hiện qua các điểm cố định lần lượt trên từng đối tượng, thay vì nhận diện số lượng một cách tổng thể [34]. Ngoài ra, trong các nhiệm vụ liên quan đến tia số, ánh nhìn của trẻ thường phân tán và thiếu chính xác, cho thấy sự hạn chế trong liên kết giữa chú ý thị giác và biểu diễn số học [35].

Trong khi đó, ở trẻ mắc chứng khó đọc, các đặc trưng chuyển động mắt phản ánh quá trình giải mã ngôn ngữ viết. Trẻ thường có thời lượng và số lượng điểm cố định cao hơn, đường trượt ngắn hơn và tần suất hồi quy lớn, cho thấy quá trình xử lý thông tin chậm và kém hiệu quả [83]. Ngoài ra, việc lặp lại điểm cố định trên cùng một từ phản ánh khó khăn trong sử dụng từ vựng và giải mã ngữ âm, trong khi kiểu chuyển động ngắt quãng làm thu hẹp phạm vi xử lý thị giác, khiến nhíp đọc rời rạc hơn. Những đặc điểm này cho thấy sự quá tải của trí nhớ làm việc và gây khó khăn trong việc tích hợp thông tin ở cấp độ câu và đoạn văn.

1.3.3 Công nghệ theo dõi chuyển động mắt trong hỗ trợ chẩn đoán

Hướng nghiên cứu ứng dụng công nghệ theo dõi chuyển động mắt trong hỗ trợ chẩn đoán trẻ rối loạn phát triển đang ghi nhận sự chuyển dịch rõ rệt từ các phân tích mang tính mô tả sang các phương pháp dự đoán tự động tích hợp học máy và học sâu [13, 30]. Thay vì chỉ dừng lại ở việc quan sát và diễn giải các đặc trưng hành vi thị giác, các nghiên cứu gần đây tập trung xây dựng các mô hình có khả năng học trực tiếp từ dữ liệu ảnh nhìn nhằm phát hiện và phân loại rối loạn một cách tự động.

Trong chẩn đoán trẻ RLPTK, các thuật toán học máy được áp dụng phổ biến để phân loại dữ liệu ảnh nhìn theo cả không gian và thời gian. Các phương pháp được áp dụng phổ biến như máy học véc tơ hỗ trợ (Support Vector Machine - SVM), mạng nơ-ron tích chập (Convolutional Neural Network - CNN) và các mô hình chuỗi như bộ nhớ dài-ngắn hạn (Long Short-Term Memory - LSTM) [84, 84]. Trong đó, SVM là một phương pháp phổ biến cho bài toán phân loại nhị phân nhờ khả năng xử lý hiệu quả dữ liệu đa chiều trong bối cảnh số lượng mẫu hạn chế, một đặc điểm thường gặp trong các nghiên cứu về RLPTK [84]. Về hiệu năng, các mô hình này đều ghi nhận tỷ lệ phân loại cao. Một nghiên cứu so sánh ba nhóm phương pháp phân loại phục vụ phát hiện sớm RLPTK [85], bao gồm: (i) các mô hình mạng nơ-ron nhân tạo và mạng truyền thẳng đạt độ chính xác lên đến 99,8%; (ii) các kiến trúc CNN như GoogleNet và ResNet-18, trong đó ResNet-18 đạt 97,6% , mức này cao hơn GoogleNet (93,6%); (iii) các mô hình lai kết hợp CNN và SVM, trong đó CNN đảm nhiệm trích xuất đặc trưng còn SVM thực hiện phân loại, đạt độ chính xác 95,5% (GoogleNet + SVM) và 94,5% (ResNet-18 + SVM). Một công trình khác tiến hành đánh giá sáu mô hình học sâu, bao gồm mạng nơ-ron hồi quy (Recurrent Neural Networks - RNN), CNN, LSTM, mô hình lai LSTM-RNN, mạng nơ-ron hồi quy với nút có cổng (Gated Recurrent Units - GRU) và bộ nhớ dài-ngắn hạn hai chiều (Bidirectional long short-term memory - BiLSTM) [85]. Kết quả chỉ ra rằng mức độ hiệu quả có sự biến động giữa các kiến trúc; trong nhóm thử nghiệm này, LSTM đạt mức 98,33%, tiếp đến là mô hình CNN-LSTM (97,94%), GRU (97,49%) và BiLSTM (96,44%).

Đối với trẻ RLTT, việc chẩn đoán sớm được triển khai thông qua các hệ thống tích hợp như hệ thống phát hiện và hỗ trợ kỹ thuật số cho học sinh

gặp khó khăn trong học tập (Digital Identification and Support of Under-Achieving Students - DIDUNAS) [86, 87]. Hệ thống này sử dụng các bài toán được số hóa trên máy tính, kết hợp với thiết bị theo dõi ánh mắt để ghi nhận chính xác chuyển động mắt của trẻ trong quá trình thực hiện nhiệm vụ. Dữ liệu thu thập sau đó được phân tích bằng các mô hình trí tuệ nhân tạo nhằm phát hiện các dấu hiệu đặc trưng của RLTT. Bên cạnh xác định nội dung mà trẻ gặp khó khăn, hệ thống còn làm rõ cách những khó khăn này biểu hiện trong quá trình nhận thức, ví dụ như sự phân tán chú ý hoặc việc sử dụng các chiến lược đếm kém hiệu quả [86, 87]. Hệ thống đã được kiểm chứng trên hơn 300 trẻ tại Đức và Cyprus, cho thấy tiềm năng trong việc triển khai các giải pháp chẩn đoán sớm tự động, góp phần hỗ trợ trẻ kịp thời và hiệu quả trên quy mô lớn.

Trong khi đó, với trẻ mắc chứng khó đọc, nhiều nghiên cứu đã chỉ ra tiềm năng của các mô hình học máy trong việc khai thác dữ liệu chuyển động mắt để phát hiện nhóm trẻ này [26]. Ví dụ, nghiên cứu của Rello [14] xây dựng mô hình SVM dựa trên các đặc trưng như số lượng điểm cố định, thời lượng điểm cố định trung bình và số lần hồi quy, đạt độ chính xác 80.18% [14]. Bên cạnh đó, thay vì sử dụng các đặc trưng đã được trích xuất như điểm cố định, đường trượt hay hồi quy, nghiên cứu của Boris Nerušil [88] tiếp cận theo hướng học trực tiếp từ dữ liệu điểm nhìn thô. Cụ thể, nhóm tác giả sử dụng tọa độ (x, y) của điểm nhìn trên màn hình làm đầu vào cho mô hình mạng nơ-ron tích chập ba lớp, với độ chính xác lên tới 96.6%.

1.3.4 Công nghệ theo dõi chuyển động mắt trong hỗ trợ can thiệp và đánh giá hiệu quả can thiệp

Mặc dù sở hữu tiềm năng ứng dụng thực tiễn, số lượng các nghiên cứu tập trung vào quy trình can thiệp và đánh giá hiệu quả thông qua công nghệ theo dõi chuyển động mắt vẫn còn hạn chế. Thực trạng này phản ánh một khoảng trống giữa lý thuyết phân tích hành vi và năng lực triển khai hệ thống trong môi trường thực tế [89]. Đối với nhóm trẻ rối loạn phát triển, các giải pháp can thiệp hiện nay chủ yếu dựa trên cơ chế phản hồi dữ liệu thị giác theo thời gian thực. Theo đó, hệ thống ghi nhận và phản hồi tức thời những thay đổi về quỹ đạo ánh nhìn, qua đó hỗ trợ trẻ duy trì trạng thái tập trung thông qua các môi trường tương tác mang tính trò chơi hóa [90, 91]. Chiến lược này bước đầu ghi nhận những kết quả tích cực trong việc gia tăng mức độ tham gia và cải thiện năng lực điều phối chú ý [90].

Tuy nhiên, để đánh giá kịp thời mức độ cải thiện nêu trên, việc chỉ dựa vào các thang đo hành vi truyền thống thường gặp nhiều rào cản do độ trễ trong biểu hiện lâm sàng. Nhằm khắc phục hạn chế này, các tham số định lượng từ chuyển động mắt, điển hình như thời lượng điểm nhìn cố định, được đề xuất làm thước đo đánh giá có độ nhạy cao [92]. Năng lực phát hiện những thay đổi nhỏ trong phân bố chú ý thị giác trong cơ chế phân bố chú ý cho phép kỹ thuật trích xuất đặc trưng thị giác cung cấp tín hiệu tiến bộ sớm, ngay cả khi trẻ chưa biểu hiện những thay đổi hành vi rõ rệt ra bên ngoài [93]. Việc áp dụng hệ thống chỉ số khách quan này cũng hỗ trợ kiểm soát những sai lệch xuất phát từ yếu tố tâm lý, điển hình là hiện tượng hiệu ứng giả dược⁴, xảy ra khi kết quả đánh giá bị chi phối bởi kỳ vọng của phụ huynh thay vì số liệu thực nghiệm [38]. Năng lực khách quan hóa dữ liệu tạo tiền đề nền tảng thiết yếu nhằm nâng cao tính minh bạch và độ tin cậy cho các báo cáo đánh giá hiệu quả can thiệp.

Trong khi đó, các nghiên cứu về ứng dụng công nghệ theo dõi chuyển động mắt trong hỗ trợ trẻ RLTT cho thấy xu hướng chuyển dịch từ chẩn đoán sang can thiệp giáo dục [95]. Một minh chứng tiêu biểu là hệ thống KI-ALF, nền tảng học toán cá nhân hóa được điều hướng bởi trí tuệ nhân tạo [96]. Khác với các mô hình phụ thuộc vào thiết bị chuyên dụng chi phí cao, KI-ALF sử dụng camera tích hợp để ước lượng và phân tích hành vi thị giác của trẻ, từ đó đánh giá năng lực tiếp thu, xác định các rào cản và đề xuất lộ trình can thiệp phù hợp [97]. Kết quả thử nghiệm tại các cơ sở giáo dục ở Đức cho thấy hệ thống không chỉ hỗ trợ hiệu quả học sinh rối loạn tính toán mà còn có giá trị ứng dụng cao trong bối cảnh thiếu hụt nguồn lực chuyên gia. Đồng thời, KI-ALF tạo nền tảng cho việc cá nhân hóa giáo dục trên diện rộng nhờ khả năng theo dõi và điều chỉnh lộ trình học tập phù hợp với từng đối tượng học sinh [97].

Trong quá trình can thiệp cho trẻ mắc chứng khó đọc, hệ thống theo dõi chuyển động mắt đảm nhiệm song song hai chức năng: một mặt là công cụ tùy chỉnh nhiệm vụ, mặt khác là thước đo đánh giá hiệu quả [98, 99]. Dựa trên dữ liệu hành vi thị giác, hệ thống tiến hành tối ưu hóa bốn thông số giao diện chính: kiểu chữ, kích thước, khoảng cách ký tự và màu sắc nền [98, 100, 101]. Việc tinh chỉnh đồng bộ các thông số này giúp khắc phục tình trạng các chữ cái chồng lấn gây nhiễu tầm nhìn, qua đó hỗ trợ

4. *Hiệu ứng giả dược* đề cập đến sự cải thiện có thể đo lường, quan sát được hoặc cảm nhận được trong tình trạng lâm sàng hoặc triệu chứng của bệnh nhân xảy ra sau khi nhận được giả dược [94].

trẻ duy trì quỹ đạo đọc liền mạch và ổn định hơn [98]. Những kết quả này cung cấp cơ sở nền tảng cho việc cải tiến tài liệu giáo dục cũng như phát triển các công nghệ hỗ trợ đọc [102]. Bên cạnh đó, các hệ thống theo dõi chuyển động mắt còn cung cấp một phương tiện định lượng khách quan để giám sát tiến độ tiếp thu của trẻ. Bằng cách đối chiếu dữ liệu hành vi thị giác trước và sau can thiệp, các chuyên gia có thể đánh giá mức độ cải thiện thông qua những thay đổi trong mô hình chuyển động mắt [32, 55]. Cụ thể, tính hiệu quả của một lộ trình can thiệp được phản ánh qua các biến thiên đặc trưng như sự suy giảm thời lượng điểm cố định, giảm tần suất hồi quy, và gia tăng biên độ của đường trượt [99, 102]. Những biến đổi định lượng này phản ánh sự thay đổi quá trình từ tiến trình giải mã ký tự mang tính tuần tự, tiêu tốn nhiều nỗ lực sang tiến trình đọc tự động [32].

1.3.5 Thảo luận

Việc tổng hợp có hệ thống các nghiên cứu hiện nay đã chỉ ra công nghệ theo dõi chuyển động mắt đóng vai trò quan trọng trong việc phân tích hành vi thị giác và hỗ trợ trẻ rối loạn phát triển. Ưu điểm của hướng tiếp cận này là khả năng cung cấp các chỉ số định lượng khách quan, hỗ trợ các phương pháp đánh giá truyền thống trong lĩnh vực giáo dục đặc biệt phụ thuộc nhiều vào quan sát chủ quan. Đặc biệt, khi kết hợp với các phương pháp học máy và học sâu, dữ liệu chuyển động mắt cho phép xây dựng các mô hình chẩn đoán tự động với hiệu năng ổn định, đồng thời mở ra khả năng cá nhân hóa can thiệp dựa trên các đặc trưng hành vi của từng trẻ.

Tuy nhiên, hiệu quả của các hệ thống này vẫn phụ thuộc vào chất lượng dữ liệu đầu vào, phương pháp trích xuất đặc trưng và thiết kế thử nghiệm. Sự khác biệt về thiết bị, cách thức thu thập và cách định nghĩa các chỉ số (điểm cố định, đường nhìn, vùng quan tâm, v.v.) dẫn đến tính không nhất quán giữa các nghiên cứu, gây khó khăn cho việc so sánh và tổng hợp kết quả. Bên cạnh đó, mặc dù các mô hình học sâu đạt hiệu năng cao, khả năng giải thích của chúng còn hạn chế, làm giảm mức độ tin cậy của các ứng dụng cả trong nghiên cứu và thực tế. Ngoài ra, còn tồn tại sự mất cân đối giữa các hướng nghiên cứu: trong khi các nghiên cứu về phân tích đặc trưng và chẩn đoán phát triển mạnh, các nghiên cứu về can thiệp và đánh giá hiệu quả can thiệp cần được triển khai thực tế vẫn còn hạn chế. Điều này cho thấy khoảng cách giữa nghiên cứu và ứng dụng, đòi hỏi sự kết hợp

chặt chẽ hơn giữa công nghệ và giáo dục đặc biệt nhằm nâng cao tính khả thi và hiệu quả trong thực tiễn.

1.3.6 Các khoảng trống nghiên cứu

Dựa trên kết quả tổng quan và hệ thống hóa các tài liệu liên quan, luận án xác định ba khoảng trống khoa học chính về phương pháp luận và kỹ thuật triển khai. Những hạn chế này được định hình thành ba bài toán trọng tâm cần giải quyết nhằm xây dựng hệ thống hỗ trợ can thiệp vòng kín:

Bài toán 1: Xây dựng bộ dữ liệu đa phương thức

Phân tích các tài liệu hiện tại cho thấy, phần lớn các nghiên cứu về hành vi thị giác thường được thực hiện trên các kích thích dạng tĩnh hoặc trong môi trường thực nghiệm bị giới hạn tương tác tự nhiên [40].

Khoảng trống nghiên cứu: Cơ sở học thuật hiện thiếu vắng các bộ dữ liệu đa phương thức có khả năng tích hợp đồng bộ tọa độ điểm nhìn, tín hiệu âm thanh và thao tác vận động đối với kích thích dạng chuỗi khung hình động. Các bộ dữ liệu trong phòng thí nghiệm thường thiếu sự tương đồng với bối cảnh học tập tự nhiên, qua đó làm giới hạn không gian đặc trưng và làm giảm khả năng khái quát của các mô hình phân tích ở giai đoạn sau. Vấn đề này được luận án giải quyết tại Chương 2.

Bài toán 2: Tự động hóa trích xuất đặc trưng không gian động và khắc phục giới hạn quy mô mẫu

Trong quy trình định lượng hành vi, nhiều nghiên cứu hiện tại vẫn phụ thuộc vào phương thức tiến hành gán nhãn thủ công hoặc đánh dấu vùng quan tâm tĩnh [46]. Đồng thời, do đặc thù y sinh, quy mô dữ liệu thu thập từ nhóm trẻ rối loạn phát triển thường hạn chế.

Khoảng trống nghiên cứu: Việc áp dụng cách xác định vùng quan tâm tĩnh lên lên các khung hình động sinh ra sai số không gian do đối tượng liên tục thay đổi hình thái và vị trí, trong khi phương pháp gán nhãn thủ công lại không đáp ứng được yêu cầu xử lý thời gian thực. Cùng với đó, môi trường học thuật hiện đang khuyết thiếu các phương pháp tăng cường dữ liệu có khả năng duy trì độ chuẩn xác sinh lý. Các phương pháp thông thường bộc lộ hạn chế trong việc bảo tồn đặc tính động lực học vi mô của nhãn cầu, điển hình như cấu trúc phổ nhiễu hồng $1/f$. Rào cản về theo dõi đối tượng động và sự thiếu hụt dữ liệu huấn luyện chất lượng này là thách thức kỹ thuật lớn sẽ được phân tích và giải quyết tại Chương 3.

Bài toán 3: Mô hình hóa nhận thức và cá nhân hóa cơ chế can thiệp

Khi xây dựng các hệ thống hỗ trợ can thiệp, xu hướng phổ biến hiện nay là ứng dụng các kiến trúc mạng học sâu để phân loại trực tiếp từ chuỗi dữ liệu.

Khoảng trống nghiên cứu: Đặc thù của lĩnh vực giáo dục đặc biệt đòi hỏi các quyết định can thiệp phải có tính lý giải rõ ràng. Tuy nhiên, các mô hình học sâu thường bộc lộ sự thiếu minh bạch trong quá trình đưa ra quyết định [48]. Các nghiên cứu hiện nay đang thiếu một kiến trúc ra quyết định lai, vừa ứng dụng tính minh bạch của hệ cơ sở luật nhằm định hướng chú ý tức thời, vừa tích hợp năng lực suy diễn xác suất của mạng Bayes động nhằm cá nhân hóa lộ trình học tập. Mục tiêu kỹ thuật này được luận án định hình và đề xuất phương pháp giải quyết tại Chương 4.

1.3.7 Giới hạn phạm vi nghiên cứu

Phạm vi nghiên cứu của luận án tập trung vào quá trình xử lý và phân tích dữ liệu hành vi thị giác kết hợp với các dữ liệu tương tác cơ bản (âm thanh, thao tác chuột) nhằm nhận diện đặc trưng và hỗ trợ định hướng can thiệp cho ba nhóm đối tượng: trẻ rối loạn phổ tự kỉ, trẻ mắc chứng khó đọc và trẻ rối loạn tính toán. Nghiên cứu giới hạn trong phạm vi tín hiệu hành vi và không bao gồm các phương pháp phân tích tín hiệu sinh lý thần kinh phức tạp như điện não đồ hoặc ảnh chụp cộng hưởng từ chức năng.

Về phương diện triển khai thực tiễn, các phương pháp và giải pháp hệ thống đề xuất được thiết kế, thử nghiệm và đánh giá chủ yếu trong môi trường thực nghiệm có kiểm soát. Do đó, luận án chưa xem xét đến các yếu tố thương mại hóa hoặc khả năng triển khai trên quy mô giáo dục đại trà, bao gồm chi phí trang bị phần cứng chuyên dụng, khả năng mở rộng hạ tầng máy chủ hay việc tích hợp trực tiếp vào các hệ thống quản lý dữ liệu giáo dục quốc gia. Ngoài ra, trong quá trình hệ thống hóa tài liệu, do phụ thuộc vào các công trình đã công bố, các kết quả tổng hợp có thể chịu ảnh hưởng bởi sự khác biệt về thiết kế thử nghiệm, độ chuẩn xác của thiết bị đo lường và tiêu chuẩn chẩn đoán giữa các nghiên cứu.

Từ những giới hạn khách quan và các khoảng trống khoa học đã phân tích, luận án xác định ba hướng nghiên cứu trọng tâm như sau:

- Xây dựng bộ dữ liệu đa phương thức: Xây dựng quy trình thu thập và đồng bộ hóa các luồng dữ liệu (chuyển động mắt, âm thanh, tương tác chuột) theo thời gian thực đối với các kích thích thị giác động. Hướng nghiên cứu này tạo ra cơ sở thực chứng để đánh giá khách quan trạng thái xử lý thông tin của trẻ.
- Đề xuất phương pháp trích xuất và tăng cường dữ liệu chuyển động mắt: Hướng nghiên cứu này phát triển một kiến trúc lai nhằm tự động nhận diện vùng quan tâm trên chuỗi khung hình động, đồng thời ứng dụng thuật toán tối ưu tiến hóa nhằm tăng cường dữ liệu quỹ đạo ánh nhìn. Giải pháp kỹ thuật này hỗ trợ cải thiện độ chính xác trong quá trình trích xuất đặc trưng và khắc phục giới hạn về quy mô mẫu huấn luyện.
- Phát triển giải pháp hỗ trợ can thiệp cá nhân hóa: Xây dựng mô hình ra quyết định lai luồng kép, kết hợp hệ cơ sở luật nhằm định hướng sự chú ý tức thời và mạng Bayes động nhằm tự động điều chỉnh lộ trình học tập phù hợp với diễn biến nhận thức của trẻ.

1.4 Các bộ dữ liệu tham chiếu và tiêu chuẩn đánh giá

1.4.1 Các bộ dữ liệu tham chiếu công khai

Mặc dù số lượng bị giới hạn bởi các rào cản về đạo đức y sinh, một số bộ dữ liệu công khai đã được cộng đồng nghiên cứu xây dựng làm cơ sở đối chuẩn, tiêu biểu bao gồm:

- Saliency4ASD: Bộ dữ liệu tập trung vào việc dự báo bản đồ phân bố sự chú ý của nhóm đối tượng rối loạn phát triển khi tiếp nhận thông tin hình ảnh tự nhiên, cung cấp cơ sở để định lượng sự khác biệt về thiên hướng tương tác không gian.
- ETDD70 (Eye Tracking Dyslexia Dataset 70): Bộ dữ liệu này bao gồm quỹ đạo tương tác của 70 học sinh, bao gồm nhóm đối tượng khó đọc và nhóm phát triển điển hình, trong quá trình thực hiện các tác vụ phân tích văn bản. Bộ dữ liệu cung cấp không gian tham chiếu để đánh giá các đặc tính nhận thức thông qua các vi chuyển động tương tác.

1.4.2 Hệ thống các độ đo đánh giá

Nhằm bảo đảm tính khách quan trong quy trình đánh giá hiệu năng của các mô hình học máy và thuật toán phân đoạn được triển khai tại Chương 2 và Chương 3, luận án áp dụng hệ thống tiêu chuẩn đo lường định lượng chuyên biệt, bao gồm:

- Độ chính xác (Accuracy): Tỷ lệ tổng số lượt dự báo chuẩn xác trên toàn bộ bộ dữ liệu kiểm thử.
- Độ nhạy (Sensitivity): Tỷ lệ nhận diện chuẩn xác các trường hợp mắc rối loạn phát triển, được tính bằng công thức $TP/(TP + FN)$. Thông số này đóng vai trò trọng yếu trong hệ thống hỗ trợ sàng lọc nhằm hạn chế tối đa việc phân loại sai nhóm đối tượng cần can thiệp lâm sàng.
- Độ đặc hiệu (Specificity): Tỷ lệ phân loại chuẩn xác các trường hợp phát triển điển hình, được tính bằng công thức $TN/(TN + FP)$.
- Chỉ số Jaccard (IoU) và hệ số Dice: Các đại lượng này được sử dụng chuyên biệt nhằm đo lường mức độ trùng khớp không gian giữa vùng phân đoạn tự động của thuật toán và nhãn đối chứng do chuyên gia xác lập.
- Diện tích dưới đường cong ROC (AUC): Chỉ số này lượng hóa năng lực phân tách tổng thể của mô hình trên nhiều ngưỡng quyết định độc lập.

Kết chương

Chương 1 đã tổng hợp các kiến thức nền tảng về công nghệ theo dõi chuyển động mắt, đồng thời phân tích các đặc trưng thị giác của nhóm trẻ rối loạn phát triển. Tiếp nối cơ sở lý thuyết này, nội dung chương tập trung hệ thống hóa các nghiên cứu liên quan, qua đó làm rõ ba khía cạnh trọng tâm: thành tựu khoa học đã đạt được, các hướng tiếp cận chủ đạo và những hạn chế hiện hữu. Hệ thống lý luận này đóng vai trò tiền đề để luận án đề xuất các giải pháp hỗ trợ trẻ rối loạn phát triển ở những chương tiếp theo. Toàn bộ kết quả tổng quan tài liệu của chương đã được công bố chính thức tại các công trình sau:

a. Các công trình đã được công bố/đăng tải

- **Thi-Quynh-Hoa Nguyen**, Thi-Kieu-Chinh Nguyen, Thanh-Ha Le, and Thi-Duyen Ngo. Eye-Tracking Technology Applications for Supporting Individuals with Autism Spectrum Disorder: Insights, Challenges, and Opportunities. In: *Disability and Rehabilitation: Assistive Technology*, 2025, 20(8), pp. 2630–2665. DOI: <https://doi.org/10.1080/17483107.2025.2531236> (WoS/SCOPUS Q1, IF: 2.1).
- **Thi-Quynh-Hoa Nguyen**, Tuan-Long Tran, Quang-Huy Trinh, and Thi-Duyen Ngo. Using Technology to Support Children with Dyscalculia in Pre-Primary and Primary School: Insights, Challenges, and Opportunities. In: *Disability and Rehabilitation: Assistive Technology*, 2026, pp. 1–24. DOI: <https://doi.org/10.1080/17483107.2026.2658104> (WoS/SCOPUS Q1, IF: 2.1).

Chương 2

XÂY DỰNG BỘ DỮ LIỆU ĐA PHƯƠNG THỨC CỦA TRẺ RỐI LOẠN PHÁT TRIỂN

Nhằm giải quyết những hạn chế về tính phân tán và thiếu đồng bộ của các nguồn dữ liệu lâm sàng đã được báo cáo trong Chương 1, Chương 2 mô tả quy trình xây dựng bộ dữ liệu đa phương thức thu thập từ trẻ rối loạn phát triển. Việc xây dựng các bộ dữ liệu tập trung vào hai nhóm trẻ: trẻ rối loạn phổ tự kỉ và trẻ rối loạn tính toán. Để bảo đảm tính nhất quán cho dữ liệu, luận án đề xuất và áp dụng khung quy trình thống nhất gồm bốn bước: thiết kế kích thích thị giác, thu dữ liệu, đồng bộ dữ liệu và tổ chức dữ liệu. Bộ dữ liệu được xây dựng từ quy trình thu nhận đa luồng này cung cấp cơ sở tham chiếu thực chứng, đóng vai trò là đầu vào trực tiếp cho các thuật toán trích xuất đặc trưng thị giác ở chương tiếp theo.

2.1 Giới thiệu

2.1.1 Vai trò của dữ liệu chuyển động mắt trong đánh giá trẻ rối loạn phát triển

Các phương pháp đánh giá truyền thống trẻ rối loạn phát triển thường dựa trên kết quả quan sát hành vi hoặc điểm số kiểm tra tĩnh, do đó chưa phản ánh đầy đủ tiến trình xử lý thông tin nội tại của trẻ [103]. Để khắc phục rào cản này, công nghệ theo dõi chuyển động mắt cung cấp một phương

pháp tiếp cận khách quan thông qua cơ chế đo lường liên tục theo thời gian thực [8, 52, 55, 72]. Khả năng thu thập luồng dữ liệu động này tạo tiền đề cho việc bóc tách các đặc điểm nhận thức đặc thù, từ đó định hướng xây dựng các mô hình đánh giá mang tính cá nhân hóa. Việc định lượng các tham số hành vi thị giác đóng vai trò là cơ sở kỹ thuật để giải quyết bài toán trên, được thể hiện qua ba khía cạnh ứng dụng:

- Đối với nhóm trẻ rối loạn phát triển mang đặc trưng khiếm khuyết tương tác xã hội: Các tham số về quỹ đạo thị giác hỗ trợ phát hiện tình trạng phân tán sự chú ý hoặc suy giảm mức độ tập trung vào vùng nhiệm vụ trọng tâm [30, 71, 73].
- Đối với nhóm trẻ rối loạn phát triển có rào cản về tư duy số học: Chỉ số thời lượng điểm nhìn cố định hỗ trợ định lượng mức độ áp lực nhận thức trong quá trình xử lý các tác vụ tính toán [10, 76, 77].
- Đối với nhóm trẻ rối loạn phát triển gặp khó khăn trong giải mã ngôn ngữ: Các chỉ số về bước nhảy nhãn cầu và tần suất nhìn lại hỗ trợ phân tách sự khiếm khuyết trong khâu nhận diện từ vựng với các rào cản về năng lực đọc hiểu [31, 55, 62].

Tổng hợp lại, luồng dữ liệu chuyển động mắt đóng vai trò là nguồn thông tin đầu vào thiết yếu. Dữ liệu này cho phép hệ thống tính toán thiết lập các lộ trình can thiệp giáo dục tương thích với từng đặc điểm phát triển riêng biệt của trẻ [13, 14, 60, 61, 84, 104].

2.1.2 Động lực và đóng góp

Mặc dù phân tích chuyển động mắt mang lại nhiều giá trị lâm sàng, các hướng nghiên cứu hiện nay vẫn gặp rào cản do sự thiếu hụt nguồn dữ liệu tham chiếu đa phương thức [104–106]. Bên cạnh đó, quy trình thu thập dữ liệu từ nhóm đối tượng trẻ em yêu cầu một môi trường thực nghiệm được kiểm soát chặt chẽ cùng các tiêu chuẩn đạo đức nghiên cứu nghiêm ngặt [89, 107, 108].

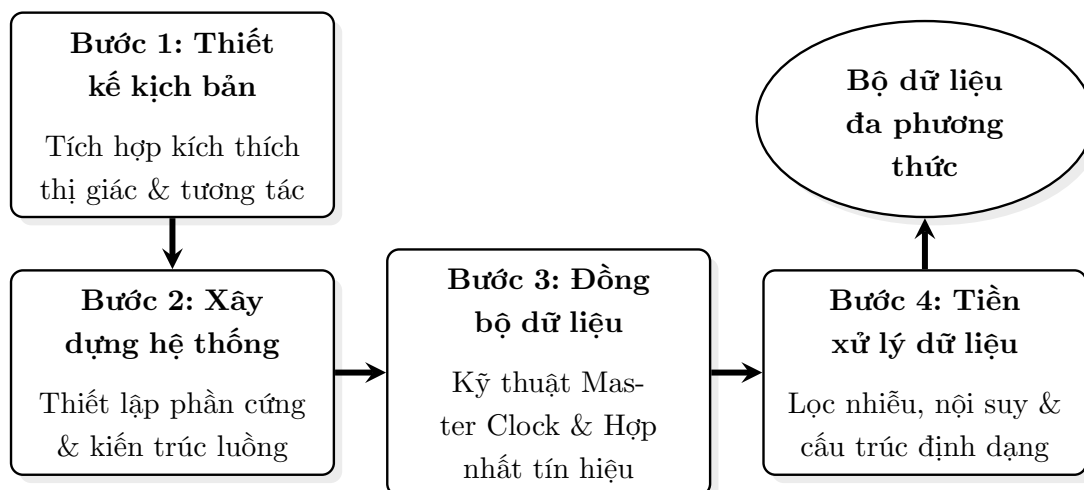
Nhằm khắc phục những rào cản nêu trên, luận án đề xuất một quy trình thống nhất trong việc thu thập dữ liệu. Bằng cách áp dụng cấu trúc đồng bộ cho các luồng thông tin, luận án đã xây dựng thành công hai bộ dữ liệu thực chứng dành riêng cho nhóm trẻ rối loạn phổ tự kỷ và trẻ rối loạn tính toán. Các bộ dữ liệu này đóng vai trò là cơ sở tham chiếu quan trọng, giúp

kiểm chứng và nâng cao độ tin cậy cho các thuật toán trích xuất đặc trưng được trình bày ở các chương tiếp theo.

2.2 Quy trình xây dựng bộ dữ liệu đa phương thức

2.2.1 Quy trình tổng quan

Để bảo đảm tính nhất quán và khả năng mở rộng khi áp dụng trên các nhóm đối tượng nghiên cứu khác nhau, luận án đề xuất quy trình xây dựng bộ dữ liệu bao gồm bốn giai đoạn. Sơ đồ luồng công việc được minh họa trong Hình 2.1. Nội dung cụ thể của từng bước được xác định như sau:



Hình 2.1: Mô hình 4 bước tổng quát xây dựng bộ dữ liệu đa phương thức

- Bước 1 (Thiết kế kịch bản thực nghiệm): Giai đoạn này tập trung thiết kế và tùy chỉnh hệ thống học liệu cùng các kích thích đa giác quan (bao gồm ảnh tĩnh, video và âm thanh) nhằm kích hoạt các phản xạ thị giác, qua đó làm bộc lộ đặc trưng nhận thức của từng nhóm đối tượng mục tiêu.
- Bước 2 (Xây dựng hệ thống thu nhận): Giai đoạn này thực hiện xây dựng kiến trúc hạ tầng phần cứng (bao gồm cảm biến theo dõi chuyển động mắt, micro và thiết bị tương tác), đồng thời thiết lập tiêu chí kiểm soát các biến số nền (như độ tuổi, giới tính và chỉ số thông minh).

- Bước 3 (Đồng bộ và hợp nhất dữ liệu): Giai đoạn này ứng dụng kỹ thuật Master Clock để xử lý độ lệch pha tần số giữa các thiết bị phần cứng [109], qua đó đồng bộ đa luồng tín hiệu về một trục thời gian thống nhất.
- Bước 4 (Tiền xử lý và kết xuất dữ liệu): Giai đoạn này tiến hành quy trình lọc nhiễu, nội suy tín hiệu và kết xuất dữ liệu thành các cấu trúc lưu trữ phân cấp (như JSON hoặc CSV). Kết quả của giai đoạn này thiết lập một nguồn dữ liệu có cấu trúc nhất quán, đóng vai trò là cơ sở thực chứng khách quan phục vụ cho các định hướng phân tích và khai thác thông tin ở những giai đoạn tiếp theo.

2.2.2 Nguyên tắc chọn mẫu và đạo đức nghiên cứu

Để triển khai quy trình 4 bước nêu trên đối với nhóm trẻ RLPTK và nhóm trẻ RLTT, luận án thiết lập các quy tắc chung về tiêu chuẩn đối tượng tham gia và đạo đức khoa học.

Quy trình tuyển chọn mẫu

Nhằm tăng độ tin cậy của dữ liệu thu thập, các đối tượng tham gia được đánh giá dựa trên hệ thống tiêu chí nhận vào và loại trừ như sau:

- Tiêu chí nhận vào:
 - Đối với nhóm trẻ RLPTK: (i) nằm trong độ tuổi từ 3 đến 7; (ii) có kết luận chẩn đoán xác định từ chuyên gia lâm sàng; (iii) có khả năng tham gia và hoàn thành nhiệm vụ thực nghiệm.
 - Đối với nhóm trẻ RLTT: (i) nằm trong độ tuổi học sinh lớp 1; (ii) đáp ứng tiêu chuẩn chẩn đoán DSM-5 về RLTT và có kết quả kiểm tra toán học ở mức thấp; (iii) có khả năng tham gia và hoàn thành nhiệm vụ thực nghiệm.
- Tiêu chí loại trừ áp dụng chung: (i) có khiếm khuyết thị giác hoặc thính giác không thể hiệu chỉnh; (ii) có rối loạn vận động hoặc thần kinh cản trở khả năng tương tác; (iii) mắc các rối loạn phát triển đồng thời khác; (iv) không hợp tác hoặc không đạt yêu cầu chất lượng khi hiệu chuẩn thiết bị.

Đạo đức nghiên cứu

Nghiên cứu đã được phê duyệt bởi Hội đồng Đạo đức Nghiên cứu thuộc Trường Đại học Sư phạm Hà Nội và tuân thủ các nguyên tắc làm việc với

đối tượng trẻ em. Trước khi thực nghiệm diễn ra, phụ huynh hoặc người giám hộ hợp pháp được cung cấp thông tin chi tiết về mục đích, quy trình, nội dung nghiên cứu và tự nguyện ký biên bản đồng thuận. Toàn bộ dữ liệu thu thập được ẩn danh và lưu trữ bảo mật, chỉ phục vụ cho mục đích nghiên cứu khoa học.

2.3 Xây dựng bộ dữ liệu hành vi thị giác của trẻ rối loạn phổ tự kỉ

2.3.1 Thiết kế kịch bản kích thích thị giác

Kịch bản thực nghiệm được thiết kế nhằm đo lường sự phân bố ánh nhìn giữa các yếu tố xã hội (khuôn mặt, biểu cảm) và các yếu tố phi xã hội (đồ vật, hình khối). Việc cân bằng hai nhóm yếu tố này cung cấp cơ sở định lượng để phân tích năng lực nhận thức của trẻ thông qua hai nhiệm vụ trọng tâm:

- Nhiệm vụ đánh giá thiên hướng chú ý giữa yếu tố xã hội và phi xã hội: Kịch bản này thiết lập các khung hình hiển thị đồng thời nút bấm cơ học (đại diện cho yếu tố phi xã hội) và khuôn mặt gấu (đại diện cho yếu tố xã hội nhân hóa) như Hình 2.2. Thiết kế thực nghiệm này nhằm kiểm chứng giả thuyết lâm sàng về xu hướng ưu tiên phân bố điểm nhìn vào các chuyển động cơ học mang tính quy luật của trẻ RLPTK, thay vì các biểu cảm mang đặc tính xã hội [110].
- Nhiệm vụ đánh giá phản ứng chú ý thị giác trước kích thích bất ngờ: Kịch bản này sử dụng video trình chiếu một đối tượng mô phỏng mang đặc tính xã hội (mô hình hộp đồ chơi chứa con rối) có sự chuyển đổi trạng thái (Hình 2.3). Việc phân chia cấu trúc video thành hai pha liên tiếp, pha tĩnh (khi hình ảnh con rối chỉ hiển thị cố định ở bốn mặt của khung hình) và pha động (khi con rối bất ngờ được bật ra từ bên trong hộp), nhằm tạo ra sự biến thiên về tính phức tạp của thông tin kích thích. Nền tảng thiết kế của hệ thống dựa trên đặc điểm nhận thức đặc thù của trẻ rối loạn phổ tự kỉ: khi đối diện với các kích thích mang yếu tố xã hội động có độ phức tạp cao, trẻ thường có phản xạ né tránh các vùng tương tác chuyển động (như vùng mắt hoặc miệng của con rối khi chuyển động bật mở) và dịch chuyển ánh nhìn về các

cấu trúc hoặc họa tiết tĩnh xung quanh nhằm tự điều chỉnh và giảm tải áp lực giác quan [111, 112].



Hình 2.2: Nhiệm vụ đánh giá thiên hướng chú ý giữa yếu tố xã hội và phi xã hội



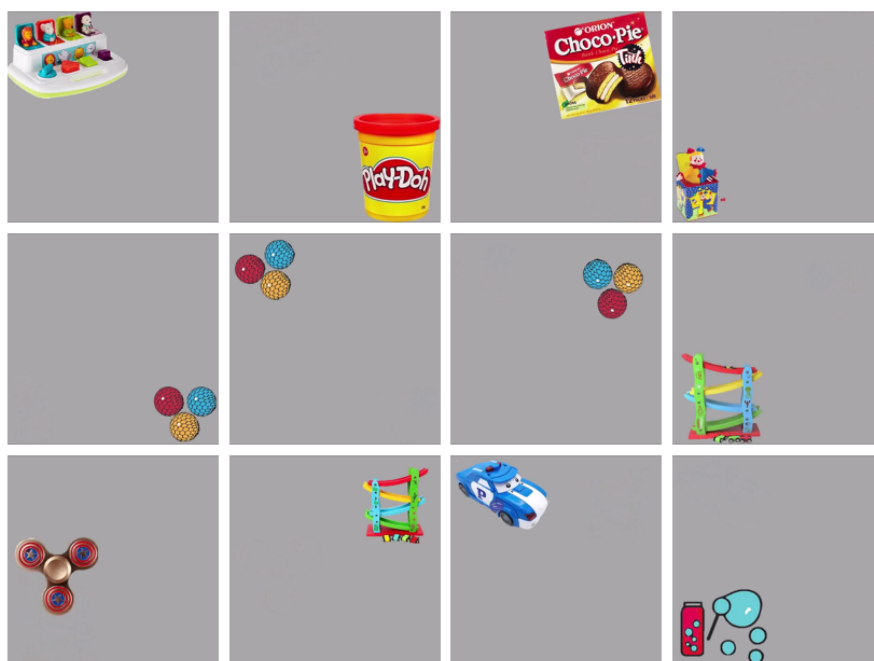
Hình 2.3: Nhiệm vụ đánh giá phản ứng chú ý thị giác trước kích thích bất ngờ

Trong mỗi phiên, số lượng hình ảnh được giới hạn tối đa 12 kích thích (mỗi kích thích hiển thị 5 giây đan xen pha định hướng 0,5 giây) nhằm giảm nguy cơ gây mệt mỏi thị giác (Hình 2.4). Phương án thiết kế này được phân tích trên hai phương diện:

- Phương diện *hiệu lực sinh thái*⁵: Các kích thích tĩnh không được lựa chọn ngẫu nhiên. Bối cảnh của 12 hình ảnh được trích xuất và thiết kế mô phỏng các tình huống tương tác thực tế trong bộ công cụ chẩn đoán lịch trình quan sát chẩn đoán tự kỉ bản thứ 2 (Autism Diagnostic Observation Schedule, Second Edition - ADOS-2) (ví dụ: một người đang cầm một đồ chơi cơ học có nút bấm) [37]. Trước khi đưa vào thực nghiệm, bộ hình ảnh này đã được thông qua hội đồng chuyên gia giáo dục đặc biệt để kiểm chứng sự phù hợp, đảm bảo phản ánh đúng các tình huống vui chơi tự nhiên hàng ngày của trẻ em Việt Nam [113].
- Phương diện kiểm soát biến số thực nghiệm: Việc sử dụng hình ảnh tĩnh trong nhiệm vụ đánh giá thiên hướng chú ý là một quyết định thiết kế thực nghiệm có chủ đích nhằm hỗ trợ tách biệt các yếu tố nhận thức. Trong môi trường kích thích động, sự dịch chuyển của vật

5. *Hiệu lực sinh thái*: Theo nhà tâm lý học phát triển Urie Bronfenbrenner, hiệu lực sinh thái là mức độ tương đồng giữa bối cảnh thực nghiệm trong các nghiên cứu khoa học và không gian sống ngoài đời thực của người tham gia [114]

thể trên khung hình có xu hướng khởi phát phản xạ hướng chú ý tự nhiên của mắt dựa trên các đặc điểm vật lý cơ bản. Bằng cách cố định khung hình, phương pháp đề xuất giảm thiểu các yếu tố nhiễu ngoại lai do chuyển động gây ra. Qua đó, quy trình thực nghiệm góp phần nâng cao độ tin cậy của phép đo, cho thấy sự phân bố và biến thiên trong quỹ đạo ánh nhìn của trẻ đối với khuôn mặt gấu hay nút bấm cơ học phản ánh trung thực hơn sự chú ý thị giác và cơ chế lựa chọn nhận thức chủ động. Việc thiết lập một không gian quan sát tĩnh giúp định hình đường cơ sở ổn định về thiên hướng chú ý nội tại, tạo nền tảng khoa học để đối sánh với các mô thức phản xạ phức tạp hơn khi trẻ tiếp xúc với kích thích động ở nhiệm vụ tiếp theo.



Hình 2.4: Video dữ liệu kích thích thị giác chứa 12 hình ảnh cho mỗi phiên

2.3.2 Thực hiện thu dữ liệu

2.3.2.1 Đối tượng tham gia

Đối tượng tham gia bao gồm 55 trẻ trong độ tuổi từ 3 đến 7. Nhóm RLPTK gồm 28 trẻ, được chẩn đoán dựa trên tiêu chuẩn DSM-5 và xác nhận bằng thang quan sát chẩn đoán RLPTK ADOS-2 [37], hai công cụ chuẩn hóa phổ biến trong đánh giá RLPTK. Nhóm không mắc RLPTK bao

gồm 27 trẻ, không có tiền sử rối loạn phát triển hoặc các vấn đề sức khỏe tâm thần. Ngoài ra, chỉ số thông minh của trẻ được đánh giá thông qua thang đo phi ngôn ngữ Leiter-3, nhằm hạn chế ảnh hưởng của yếu tố ngôn ngữ [115]. Các công cụ này hỗ trợ duy trì tính nhất quán trong việc xác định đặc điểm nhận thức và lâm sàng của đối tượng tham gia. Bảng 2.1 trình bày về đặc trưng nhân khẩu học và lâm sàng của mẫu nghiên cứu.

Bảng 2.1: Đặc điểm nhân khẩu học và lâm sàng của trẻ tham gia

Đặc điểm	RLPTK (n=28)	Không RLPTK (n=27)
Tuổi	3-7	3-7
Giới tính (Nam/Nữ)	23/5	16/11
IQ	~ 60	Bình thường
Điểm ADOS-2	Đáp ứng DSM-5	Không áp dụng

2.3.2.2 Quy trình triển khai

Quy trình thu thập dữ liệu được triển khai theo ba giai đoạn chính: hiệu chuẩn hệ thống, xác thực chất lượng dữ liệu ánh nhìn và thực hiện kịch bản trình chiếu kích thích.

- Giai đoạn 1 - Hiệu chuẩn cá nhân hóa: Xuất phát từ những rào cản về năng lực duy trì sự chú ý trực quan và khó khăn trong việc cố định ánh nhìn vào các mục tiêu tĩnh ở nhóm trẻ rối loạn phát triển, việc áp dụng phương pháp hiệu chuẩn tiêu chuẩn thường gặp nhiều hạn chế và khó đảm bảo độ chính xác. Nhằm khắc phục rào cản này, luận án triển khai giao thức hiệu chuẩn cá nhân hóa bao gồm 5 điểm.
- Giai đoạn 2 - Xác thực chất lượng: Sau khi hiệu chuẩn, hệ thống thực hiện bước xác thực nhằm đánh giá độ chính xác không gian của dữ liệu ánh nhìn. Hệ thống hiển thị ngẫu nhiên 4 điểm kiểm tra và sai số được tính bằng khoảng cách Euclidean giữa tọa độ điểm nhìn dự đoán (x_{gaze}, y_{gaze}) và vị trí mục tiêu (x_{target}, y_{target}) .

Ngưỡng chấp nhận được thiết lập ở mức $\leq 1.5^\circ$ góc nhìn, tương đương khoảng 45 pixels với khoảng cách quan sát 60cm [51]. Các phiên đo vượt quá ngưỡng sai số này sẽ bị loại và tiến hành hiệu chuẩn lại tối đa 3 lần. Nếu không đạt yêu cầu sau 3 lần thử, đối tượng sẽ bị loại khỏi phân tích nhằm duy trì tính toàn vẹn của dữ liệu [113].

- Giai đoạn 3 - Trình chiếu kích thích: Sau khi qua bước xác thực, trẻ tham gia thực hiện kịch bản thực nghiệm được thiết kế theo chu kỳ ba pha lặp lại:
 - Pha định hướng: Hiện thị điểm đánh dấu trung tâm trong 0,5 giây nhằm thu hút ánh nhìn.
 - Pha kích thích: Trình chiếu hình ảnh mục tiêu trong 5 giây để thu thập dữ liệu hành vi thị giác.
 - Pha tái định hướng: Lặp lại điểm đánh dấu trong 0,5 giây để chuẩn bị cho chu kỳ tiếp theo.

2.3.3 Tiền xử lý và đồng bộ hóa dữ liệu

Quy trình tiền xử lý được thực hiện qua các bước sau:

- Đồng bộ hóa trực thời gian: Các tín hiệu chuyển động mắt (thu nhận với tần số danh định⁶ 90Hz) được ánh xạ về trục thời gian chuẩn với chu kỳ lấy mẫu 1ms (1000Hz). Các nhãn thời gian được đồng bộ hóa với kịch bản hệ thống trình chiếu để xác định mốc thời gian bắt đầu (t_{start}) và kết thúc (t_{end}) của từng pha kích thích ảnh tĩnh (5 giây).
- Phân tích suy hao tín hiệu và cơ chế xử lý khoảng trống: Trong quá trình đo lường, chuỗi tín hiệu có thể bị ngắt quãng do thiết bị mất tín hiệu theo dõi chuyển động mắt (ví dụ: trẻ chớp mắt hoặc quay đầu). Hệ thống áp dụng cơ chế xử lý dữ liệu khuyết dựa trên ngưỡng thời gian:
 - *Khoảng khuyết tín hiệu ngắn* ($Gaps \leq 300ms$): Chủ yếu do chớp mắt sinh lý. Thuật toán nội suy tuyến tính được áp dụng dựa trên tọa độ hai điểm biên để điền khuyết chuỗi tín hiệu trên trục cơ sở.
 - *Khoảng khuyết tín hiệu dài* ($Gaps > 300ms$): Xảy ra khi trẻ quay hẳn đầu khỏi màn hình hoặc thiết bị mất tín hiệu theo dõi kéo dài. Hệ thống không nội suy để tránh làm sai lệch dữ liệu hành vi. Các khoảng này được gán giá trị rỗng và chuỗi ảnh nhìn tự động được phân tách thành các phân đoạn xử lý độc lập.

6. *Tần số danh định* (Nominal Sampling Rate): là thông số lấy mẫu lý thuyết do nhà sản xuất thiết bị công bố. Trong quá trình vận hành thực tế, tần số này thường xuyên dao động do độ trễ của hệ điều hành và tình trạng rớt khung hình. Theo Holmqvist và cộng sự (2011) cũng như Duchowski (2017), sự sai lệch ngẫu nhiên này đòi hỏi một bước đồng bộ hóa bắt buộc để ánh xạ tín hiệu về trục thời gian chuẩn [51, 116]

- Ánh xạ không gian và gắn kết sự kiện: Dựa trên hệ trục thời gian đã đồng bộ, tọa độ điểm nhìn (x, y) được ánh xạ trực tiếp vào lưới tọa độ không gian của các vùng quan tâm trên khung hình tương ứng. Nếu tổng thời lượng các khoảng khuyết tín hiệu dài (Gaps > 300ms) vượt quá 30% thời lượng của một phiên đo, dữ liệu của phiên đó sẽ tự động bị loại bỏ nhằm duy trì độ tin cậy của bộ dữ liệu.

Quy trình thiết lập nhãn dữ liệu đối chứng: Nhằm xây dựng bộ dữ liệu tham chiếu phục vụ quá trình huấn luyện và kiểm định thuật toán tại Chương 3, thao tác gắn nhãn ranh giới vùng quan tâm được triển khai thông qua nền tảng công cụ chú thích thị giác máy tính (Computer Vision Annotation Tool - CVAT). Việc xác định ranh giới không gian được các chuyên gia giáo dục đặc biệt tiến hành một cách độc lập. Kết quả gắn nhãn chỉ được tích hợp vào bộ dữ liệu khi hệ số đồng thuận liên quan sát viên đạt ngưỡng quy định (Cohen's $\kappa > 0,85$). Cơ chế kiểm duyệt này được thiết lập nhằm kiểm soát phương sai và giảm thiểu sai số thiên lệch trong quá trình đánh giá hệ thống.

2.3.4 Tổ chức lưu trữ dữ liệu

Về mặt tổ chức lưu trữ, dữ liệu được phân chia thành các thư mục riêng biệt tương ứng với từng đối tượng tham gia. Đối với mỗi phiên thực nghiệm, dữ liệu được hệ thống kết xuất thành hai tệp: tệp `ET.json` lưu trữ các tọa độ điểm nhìn thô theo định dạng phân cấp và tệp `sync_data.csv` tập hợp 14 trường đặc trưng động học (bao gồm các chỉ số tiêu biểu như `Gaze_Velocity`, `Is_Fixation`, `Fixation_ID`). Chi tiết về bộ dữ liệu được trình bày tại Bảng 2.2.

Bảng 2.2: Bộ dữ liệu chuyển động mắt của trẻ rối loạn phổ tự kỉ

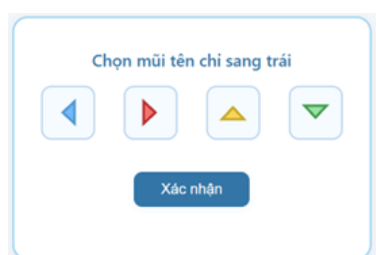
Đặc điểm	Một phiên	Toàn bộ bộ dữ liệu (660 phiên)
Thời lượng kích thích thị giác	~78 giây (1,3 phút)	~51.480 giây (~14,3 giờ)
Tần số lấy mẫu	90 Hz	90 Hz
Số dòng dữ liệu (điểm nhìn)	~7.020 dòng	~4.633.200 dòng (~4,6 triệu)
Số điểm cố định trung bình	~200 điểm cố định	~132.000 điểm cố định
Ngưỡng dữ liệu hợp lệ (tỷ lệ quét)	$\geq 70\%$	$\geq 70\%$

2.4 Xây dựng bộ dữ liệu đa phương thức của trẻ rối loạn tính toán

2.4.1 Thiết kế kịch bản học liệu tương tác

Dựa trên lý thuyết nhận thức toán học [24, 117], kịch bản cho trẻ RLTT được phân loại theo các nhóm năng lực:

- Xử lý không gian thị giác: Phân biệt hướng (Hình 2.5) và xác định vị trí tương đối.
- Ước lượng độ lớn (trục số): Định vị tọa độ trên trục số một chiều nhằm đánh giá khả năng thiết lập biểu diễn không gian số trong tâm trí [118] (Hình 2.6).
- Nhận biết số lượng nhanh: Nhận diện tức thời nhóm số lượng nhỏ không qua thao tác đếm (Hình 2.7).
- Phối hợp thị giác - thính giác: Thông tin nhiệm vụ phát qua giọng nói, yêu cầu lựa chọn đáp án hiển thị bằng ký hiệu chữ số (Hình 2.8).



Hình 2.5: Nhiệm vụ không gian thị giác



Hình 2.6: Nhiệm vụ định vị trục số



Hình 2.7: Nhận biết số lượng nhanh



Hình 2.8: Phối hợp thị giác - thính giác

2.4.2 Thực hiện thu nhận đa luồng tín hiệu

2.4.2.1 Đối tượng tham gia

Đối tượng tham gia bao gồm 41 học sinh lớp 1 với độ tuổi trung bình 6.0 ± 0.3 . Nhóm trẻ RLTT gồm 20 em, được xác định dựa trên tiêu chuẩn chẩn đoán của DSM-5 đối với rối loạn học tập đặc hiệu trong lĩnh vực toán học [9]. Việc sàng lọc được thực hiện qua kết quả các bài kiểm tra năng lực toán học chuẩn hóa, trong đó các trẻ thuộc nhóm này có điểm số nằm trong 30% thấp nhất, đồng thời được đánh giá chéo bởi giáo viên chủ nhiệm và chuyên gia tâm lý qua các bài kiểm tra nhận thức số lượng cơ bản. Nhóm không mắc RLTT gồm 21 trẻ, không có biểu hiện khó khăn trong học tập và không có tiền sử rối loạn phát triển.

Chỉ số thông minh của đối tượng tham gia được đánh giá bằng thang đo Raven [119], đảm bảo các trẻ đều có mức IQ trong khoảng bình thường hoặc cao ($IQ > 85$) [120], qua đó loại trừ ảnh hưởng của khuyết tật trí tuệ tổng quát. Ngoài ra, các trẻ đều được các chuyên gia y tế đánh giá thị lực và thính lực bình thường (hoặc đã được hiệu chuẩn đạt mức bình thường), không đồng mắc các rối loạn thần kinh như tăng động giảm chú ý hoặc RLPTK [121].

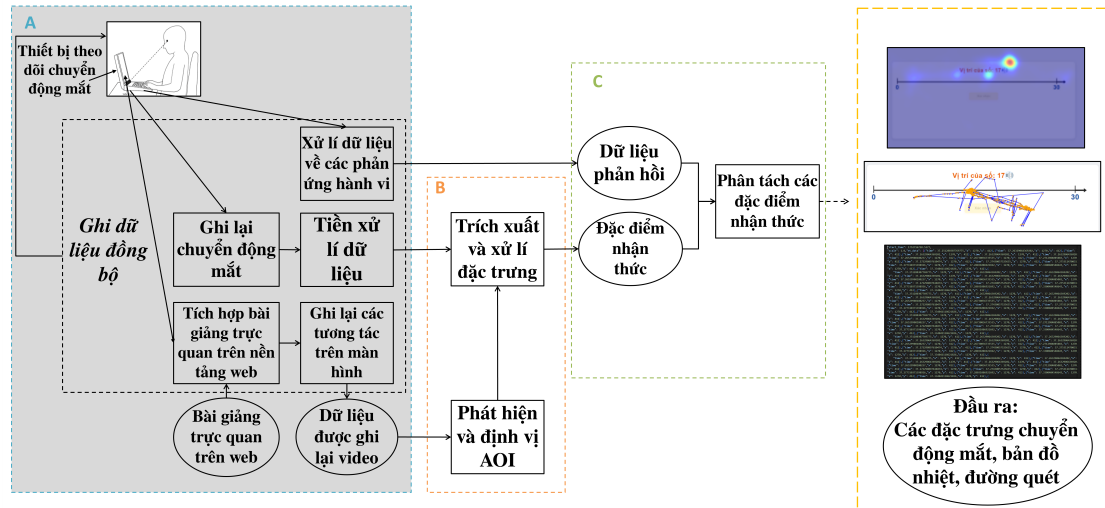
Bên cạnh đó, nhằm kiểm soát các yếu tố môi trường, các trẻ sinh sống tại khu vực thành thị và xuất thân từ các gia đình có điều kiện kinh tế - xã hội ở mức trung bình trở lên [122]. Điều này hướng tới các em đã được tiếp cận đầy đủ với môi trường giáo dục mầm non, qua đó hạn chế ảnh hưởng của các yếu tố ngoại sinh đến kết quả nghiên cứu. Bảng 2.3 thống kê các đặc điểm nhân khẩu học và lâm sàng của mẫu nghiên cứu.

Bảng 2.3: Đặc điểm nhân khẩu học và lâm sàng của trẻ tham gia

Đặc điểm	Rối loạn tính toán (n=20)	Đối chứng (n=21)
Tuổi (Mean $\pm$ SD)	6.0 ± 0.3	6.0 ± 0.3
Giới tính (Nam/Nữ)	13/7	12/9
IQ (Raven)	> 85	> 85
Điểm kiểm tra toán	30% thấp nhất	70% còn lại

2.4.2.2 Hệ thống thu nhận

Hệ thống thu nhận dữ liệu (Hình 2.9) được thiết kế theo kiến trúc đa phương thức. Hệ thống tích hợp ba luồng dữ liệu song song nhằm ghi nhận đồng bộ các hành vi tương tác của trẻ trong quá trình thực hiện nhiệm vụ. Mỗi luồng dữ liệu được sử dụng để phân tích độc lập nhưng bổ trợ chặt chẽ cho nhau trong kiến trúc chung.



Hình 2.9: Hệ thống thu thập dữ liệu đa phương thức cho trẻ rối loạn tính toán

- Luồng dữ liệu điểm nhìn: Hệ thống sử dụng cảm biến Tobii 4C (tần số lấy mẫu 90Hz) kết hợp màn hình hiển thị 13 inch (60Hz) để ghi nhận liên tục quỹ đạo ánh nhìn thô dưới dạng chuỗi tọa độ theo thời gian (x, y, t) . Dữ liệu này phản ánh trực tiếp quá trình phân bổ chú ý thị giác của trẻ.
- Luồng dữ liệu âm thanh: Hệ thống thu thập toàn bộ tín hiệu âm thanh phản hồi của trẻ thông qua micro đơn hướng (16.000 Hz, 16-bit PCM) và lưu trữ dưới định dạng .wav. Luồng dữ liệu này cung cấp hai lớp thông tin cơ bản:
 - Nội dung ngôn ngữ: Hệ thống áp dụng thuật toán nhận dạng giọng nói để xử lý tín hiệu âm thanh thô nhằm trích xuất ngữ nghĩa, phục vụ đánh giá tính chính xác của câu trả lời toán học.
 - Đặc trưng thời gian: Hệ thống áp dụng cơ chế phát hiện hoạt động giọng nói dựa trên ngưỡng năng lượng để xác định các đoạn phát âm hợp lệ, từ đó tự động sinh ra nhãn thời gian Audio_Active.

Các nhân này đóng vai trò cơ sở để đồng bộ hóa luồng dữ liệu âm thanh với các luồng dữ liệu khác.

- Luồng dữ liệu thao tác vận động tay: Hệ thống tiến hành ghi nhận quỹ đạo di chuyển liên tục và các sự kiện thao tác chuột tại tọa độ không gian (x, y) trên mặt phẳng giao diện. Các sự kiện này là những điểm tương tác rời rạc phát sinh khi trẻ thực hiện hành động nhấp chuột, giữ hoặc kéo thả. Dữ liệu từ luồng này đóng vai trò phản ánh khách quan hành vi tương tác và phương thức điều khiển trực tiếp của trẻ đối với giao diện hệ thống.

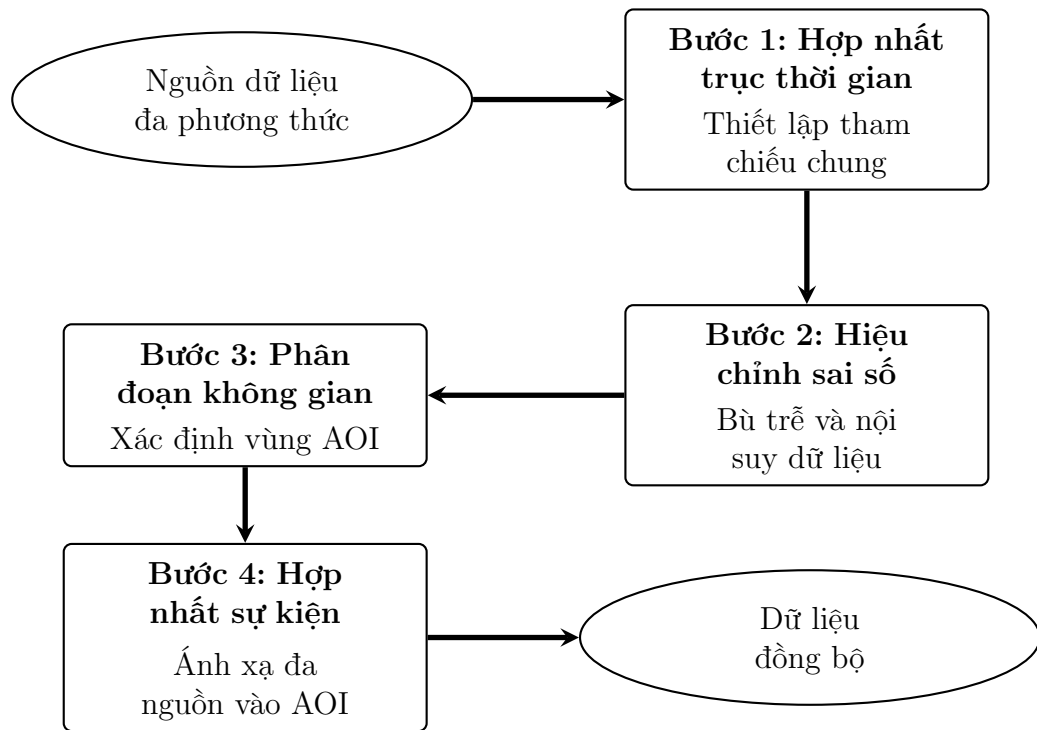
Việc tích hợp đồng bộ ba luồng dữ liệu này giúp hệ thống thiết lập mô hình biểu diễn đa chiều về tiến trình xử lý nhận thức. Cụ thể, luồng dữ liệu điểm nhìn phản ánh trạng thái chú ý thị giác, luồng dữ liệu âm thanh cung cấp thông tin về tốc độ đáp ứng và khả năng xử lý ngôn ngữ, trong khi luồng dữ liệu thao tác vận động ghi nhận các quyết định tương tác vật lý của người học. Sự hội tụ thông tin từ ba luồng này tạo cơ sở để mô hình học máy tiến hành phân tích chuyên sâu mối tương quan giữa mức độ tập trung, tốc độ xử lý thông tin và hành vi phản hồi thực tế của trẻ.

2.4.3 Đồng bộ và hợp nhất dữ liệu đa phương thức

Để xử lý sự bất đồng bộ giữa các luồng dữ liệu thu nhận từ nhiều thiết bị, luận án đề xuất quy trình đồng bộ hóa đa nguồn gồm bốn bước xử lý tuần tự (Hình 2.10).

Chi tiết kỹ thuật từng bước như sau:

- Bước 1 (Hợp nhất trục thời gian): Hệ thống thiết lập một trục thời gian cơ sở với độ phân giải 1ms (tương đương tần số 1000Hz). Thang đo 1ms này đóng vai trò là lưới tham chiếu chung, cho phép quy chuẩn các dấu thời gian dạng số thập phân từ cảm biến mắt (chu kỳ xấp xỉ 11,11ms) và luồng video (chu kỳ lớn hơn 33ms) về cùng một hệ trục tọa độ nguyên thống nhất, đồng thời bảo toàn tính nguyên vẹn của dữ liệu gốc.
- Bước 2 (Hiệu chỉnh sai số và nội suy tần số): Do giới hạn về tài nguyên xử lý của hệ thống trong quá trình thu nhận đồng thời đa luồng dữ liệu, tần số lấy mẫu thực tế của thiết bị theo dõi chuyển động mắt ghi nhận sự sụt giảm xuống mức trung bình 68,7Hz (so với tần số danh định 90Hz). Giai đoạn này tiến hành tính toán và bù trừ độ lệch



Hình 2.10: Quy trình bốn bước đồng bộ hóa dữ liệu đa nguồn

pha giữa các luồng với ngưỡng $\leq 15\text{ms}$. Đồng thời, thuật toán nội suy tuyến tính được áp dụng để điền khuyết cho các khoảng mất tín hiệu vi mô ($\leq 300\text{ms}$) do hiện tượng chớp mắt sinh lý của trẻ.

- Bước 3 (Phân tách khung hình và xác định vùng quan tâm): Hệ thống phân tách luồng dữ liệu video thành các khung hình độc lập và định lượng thời gian tồn tại thực tế của từng khung hình dựa trên cơ chế Master Clock. Tại mỗi khung hình, hệ thống xác lập lưới tọa độ không gian cho các vùng quan tâm tương ứng.
- Bước 4 (Hợp nhất sự kiện đa nguồn): Do chu kỳ hiển thị của một khung hình video thực tế kéo dài hơn chu kỳ lấy mẫu của mắt, hệ thống áp dụng cơ chế đối khớp không gian - thời gian theo cấu trúc "nhiều - một". Nhiều điểm mẫu tọa độ ánh nhìn rơi vào cùng một khoảng thời gian hiển thị sẽ được nhóm và ánh xạ vào khung hình và cấu trúc AOI tương ứng tại mốc thời gian đó. Song song đó, các thao tác điều khiển chuột và nhấn thời gian thính giác Audio_Active (khoảng $[t_{start}, t_{start} + duration]$) được tích hợp trực tiếp, liên kết cố định với mã định danh của tệp âm thanh .wav tương ứng. Chu trình kết thúc bằng một tiêu chí kiểm soát chất lượng: nếu tỷ lệ mất tín hiệu thính giác kéo dài ($> 300\text{ms}$) vượt quá 30% tổng thời lượng phiên tương

tác, toàn bộ dữ liệu đồng bộ của phiên đó sẽ được loại khỏi tập phân tích nhằm bảo toàn tính nhất quán của bộ dữ liệu tham chiếu.

2.4.4 Tổ chức lưu trữ dữ liệu

Về mặt tổ chức lưu trữ, dữ liệu được phân chia thành các thư mục riêng biệt tương ứng với từng đối tượng tham gia. Đối với mỗi phiên thực nghiệm, toàn bộ dữ liệu được hệ thống kết xuất thành nhóm bốn tài liệu độc lập: tệp ET.json, tệp Mouse.json, tệp âm thanh nguyên bản và tệp dữ liệu tổng hợp complete_synchronized_data.csv đóng vai trò lưu trữ 17 trường đặc trưng cơ sở.

Khác với luồng tọa độ mắt mang tính liên tục, các dữ liệu thao tác chuột và tín hiệu âm thanh có đặc tính rời rạc. Do đó, hệ thống tiến hành thu thập các luồng thông tin này dựa trên cơ chế hướng sự kiện (bao gồm các hành vi nhấp chuột, kéo thả hoặc phát âm). Ngay sau khi ghi nhận, hệ thống thực hiện đồng bộ trực tiếp các sự kiện rời rạc này vào cùng một trục thời gian với luồng dữ liệu điểm nhìn. Chi tiết về bộ dữ liệu cho trẻ RLTT được trình bày cụ thể tại Bảng 2.4.

Bảng 2.4: Thống kê bộ dữ liệu đa phương thức của trẻ rối loạn tính toán

Đặc điểm	Một phiên	Toàn bộ bộ dữ liệu (82 phiên)
Thời lượng phiên	342 giây (~5,7 phút)	~28.000 giây (~7,8 giờ)
Tần số lấy mẫu mắt trung bình	68,7 Hz	68,7 Hz
Số dòng dữ liệu (đồng bộ)	23.508 dòng	~1.930.000 dòng
Kích thước file CSV đồng bộ	2,67 MB	~219 MB
Độ phủ audio / dữ liệu chuột	3,1% / 0,1%	Tương tự

2.5 Thảo luận về cơ sở khoa học và kiểm soát rủi ro

Để đánh giá mức độ tin cậy và giá trị thực tiễn của các bộ dữ liệu đa phương thức đã thu thập, phần này sẽ phân tích chi tiết về cơ sở thiết kế

cũng như các biện pháp kiểm soát rủi ro trong suốt quá trình thực nghiệm.

2.5.1 Quy mô mẫu và năng lực khái quát hóa của dữ liệu sinh trắc học

Một trong những đặc điểm của *dữ liệu sinh trắc học thị giác*⁷ là độ tin cậy và giá trị thống kê không chỉ phụ thuộc vào số lượng đối tượng tham gia ($N = 55$ ở nhóm RLPTK và $N = 41$ ở nhóm RLTT), mà còn được quyết định bởi mật độ lấy mẫu trên từng cá nhân. Với thiết bị theo dõi chuyển động mắt hoạt động ở tần số 90Hz, mỗi đối tượng cung cấp một chuỗi thời gian chứa hàng chục nghìn điểm tọa độ. Để kiểm soát nguy cơ quá khớp khi huấn luyện các mô hình học máy ở giai đoạn sau, chiến lược kiểm chứng chéo theo cấp độ đối tượng (Subject-wise Leave-One-Out Cross-Validation, LOOCV) [123] được áp dụng. Cơ chế này đảm bảo hệ thống có khả năng khái quát hóa các đặc trưng rối loạn cốt lõi, qua đó ngăn ngừa tình trạng mô hình ghi nhớ cục bộ các mẫu hành vi cá biệt của từng đối tượng.

2.5.2 Đặc tính phân phối thưa của dữ liệu hướng sự kiện

Số liệu thống kê tại Bảng 2.4 cho thấy tỷ trọng của luồng dữ liệu âm thanh (3,1%) và luồng dữ liệu thao tác chuột (0,1%) thấp hơn đáng kể so với luồng dữ liệu điểm nhìn. Sự chênh lệch này xuất phát từ bản chất hướng sự kiện của luồng dữ liệu âm thanh và thao tác chuột, khác biệt với tính liên tục của luồng dữ liệu điểm nhìn. Trong các tiến trình tương tác được ghi nhận, các sự kiện thính giác và vận động phát sinh không thường xuyên tại các thời điểm rời rạc.

Hệ thống khai thác các sự kiện phân phối thưa này làm mốc thời gian tham chiếu để phân đoạn luồng dữ liệu liên tục. Cụ thể:

- Sự kiện âm thanh (Audio_Active) xác định mốc thời gian T_{start} , biểu thị thời điểm hoàn tất quá trình tiếp nhận thông tin thính giác.
- Sự kiện tương tác chuột (Mouse_Click) xác định mốc thời gian T_{end} , biểu thị thời điểm thực hiện quyết định vận động.

7. *Dữ liệu sinh trắc học thị giác*: Dữ liệu sinh trắc học thị giác là các đo lường định lượng về cách thức cá nhân tiếp nhận và xử lý thông tin thông qua hành vi chuyển động mắt. Khác với các dữ liệu sinh trắc học tĩnh (vân tay, mống mắt) dùng để nhận dạng, dữ liệu này phản ánh động lực học của sự chú ý và các trạng thái nhận thức ẩn của trẻ [124].

Hai mốc thời gian này xác định giới hạn thời gian để phân đoạn chuỗi quỹ đạo điểm nhìn, hỗ trợ tính toán khoảng thời gian xử lý thông tin hình ảnh của trẻ (độ trễ nhận thức).

2.5.3 Đặc điểm của dữ liệu đặc trưng đa phương thức

Sự tích hợp đồng bộ các luồng dữ liệu điểm nhìn, âm thanh và thao tác chuột tạo ra một tập hợp đặc trưng đa chiều, giúp khắc phục các giới hạn của phương pháp phân tích đơn phương thức. Trong đó, ưu thế của cách tiếp cận này ở khả năng phân tách rõ ràng các giai đoạn xử lý nhận thức.

Nếu chỉ sử dụng riêng lẻ luồng dữ liệu điểm nhìn, hệ thống có thể thiếu cơ sở để xác định chính xác mốc thời gian trẻ bắt đầu quá trình tư duy sau khi tiếp nhận yêu cầu bài tập. Sự thiếu hụt này trực tiếp gây ra sai số trong quá trình ước lượng tốc độ nhận thức. Việc đồng bộ dữ liệu đa phương thức giải quyết vấn đề này thông qua việc phân định rõ ràng giữa giai đoạn tiếp nhận thông tin thính giác và giai đoạn xử lý thông tin trực quan. Khả năng phân tách này đóng vai trò nền tảng giúp tăng độ chính xác cho các thuật toán đánh giá và sàng lọc đặc trưng rối loạn phát triển ở giai đoạn sau.

2.5.4 Kiểm soát dữ liệu hao hụt và rủi ro thiên lệch

Nhằm bảo vệ tính toàn vẹn của dữ liệu, quy trình sàng lọc áp dụng các ngưỡng kỹ thuật đối với sai số hiệu chuẩn, tỷ lệ quỹ đạo hợp lệ và khoảng trống tín hiệu. Quá trình này dẫn đến tỷ lệ hao hụt dữ liệu dao động từ 15% đến 18% trên tổng mẫu nghiên cứu ban đầu gồm 65 trẻ RLPTK và 50 trẻ RLTT. Tỷ lệ hao hụt này nằm trong biên độ cho phép của các nghiên cứu sử dụng thiết bị theo dõi chuyển động mắt đối với đối tượng trẻ em [76].

Nhằm kiểm soát rủi ro thiên lệch hệ thống, luận án tiến hành một kiểm định độc lập để đối sánh các đặc tính nền, bao gồm điểm chuẩn đoán ADOS-2 và chỉ số IQ trung bình, giữa nhóm dữ liệu được bảo lưu và nhóm dữ liệu hao hụt. Kết quả kiểm định ($p > 0,05$) cho thấy không tồn tại sự khác biệt mang ý nghĩa thống kê giữa hai nhóm. Phân tích này chỉ ra rằng tình trạng khuyết thiếu dữ liệu diễn ra mang tính ngẫu nhiên, chủ yếu bắt nguồn từ các yếu tố sinh lý hoặc điều kiện ngoại cảnh nhất thời như tình trạng mất mát hay sai tư thế, qua đó bảo toàn tính đại diện của mẫu nghiên cứu.

Các số liệu định lượng về hiệu năng của việc kết hợp đa phương thức dữ liệu sẽ được trình bày thông qua các thực nghiệm đối chứng tại Chương 3 và phần đánh giá kiến trúc mạng Bayes động tại Chương 4.

2.6 Đánh giá độ tin cậy của bộ dữ liệu

Nhằm kiểm chứng tính hợp lệ và năng lực phân tách của nguồn dữ liệu nguyên bản trước khi triển khai các thuật toán trích xuất, luận án tiến hành thiết lập một thử nghiệm cơ sở. Thử nghiệm này tận dụng các đặc trưng thống kê hình học cơ bản, được trích xuất trực tiếp từ chuỗi tọa độ điểm nhìn thô mà không áp dụng cơ chế nhận diện vùng quan tâm tự động, sau đó tích hợp vào mô hình phân loại SVM sử dụng nhân RBF.

Quá trình kiểm thử được thực thi trên tập mẫu bao gồm 55 đối tượng, trong đó có 28 trẻ rối loạn phát triển và 27 trẻ phát triển điển hình, thông qua chiến lược kiểm chứng chéo luân phiên từng cá thể. Kết quả thực nghiệm ban đầu ghi nhận hệ thống đạt độ chính xác 58,2% cùng chỉ số AUC ở mức 0,545. Mặc dù mức hiệu năng phân loại này chưa đáp ứng tiêu chuẩn đối với các ứng dụng sàng lọc thực tiễn, các chỉ số trên vẫn phản ánh tín hiệu phân tách cơ sở giữa hai nhóm đối tượng. Yếu tố này đóng vai trò thiết lập nền tảng dữ liệu cần thiết nhằm phát triển và xác thực phương pháp lai tối ưu hóa đặc trưng không gian và thời gian tại Chương 3.

Kết chương

Chương này đã trình bày quy trình bốn bước nhằm xây dựng hai bộ dữ liệu đa phương thức cho hai nhóm đối tượng bao gồm trẻ RLPTK và trẻ RLTT. Để phù hợp với năng lực nhận thức đặc thù của từng nhóm, luận án tiến hành thiết kế phân cấp các kịch bản tương tác, đồng thời cấu hình hệ thống nhằm chuyển đổi từ phương thức thu thập đơn luồng sang kiến trúc đồng bộ đa luồng dữ liệu.

Quá trình thực nghiệm được triển khai dựa trên các tiêu chí phân loại đối tượng, kết hợp với cơ chế xử lý hướng sự kiện và quy trình kiểm soát sai số đồng bộ thời gian dưới mức 15 ms. Cơ chế xử lý này hỗ trợ giảm thiểu nhiễu do cử động vật lý hoặc thay đổi môi trường ngoài, đồng thời duy trì tính nhất quán cho chuỗi thông tin đầu vào.

Bộ dữ liệu thu nhận đóng vai trò là cơ sở để thực hiện trích xuất đặc trưng thị giác, phục vụ quá trình huấn luyện mô hình học máy nhằm phân loại nhóm đối tượng rối loạn phát triển tại Chương 3, đồng thời cung cấp các biến quan sát thực chứng để xây dựng cơ sở luật hỗ trợ can thiệp tại Chương 4. Các đóng góp khoa học liên quan đến quy trình thiết lập dữ liệu trong chương này đã được công bố chính thức tại các công trình sau:

a. Các công trình đã được công bố/đăng tải

- **Thi-Quynh-Hoa Nguyen**, Duy-Hai Nguyen, Thanh-Ha Le, and Thi-Duyen Ngo. An Eye-Tracking System for Extracting and Visualizing Visual Features of Dyscalculia in Children. In: *Proceedings of the 14th International Symposium on Information and Communication Technology (SoICT 2025)*, Nha Trang, Vietnam, 2025. DOI: https://doi.org/10.1007/978-981-92-2590-3_31

b. Các công trình đã nộp (Đang trong quá trình xét duyệt)

- **Thi-Quynh-Hoa Nguyen**, Nu-Tam-An Nguyen, Thi-Cam-Huong Nguyen, Thanh-Ha Le, and Thi-Duyen Ngo. An automated eye-tracking framework: From semantic gaze mapping to human-in-the-loop optimization. Submitted to *Autism Research*, 2026 (**WoS/SCOPUS Q1**).

Chương 3

PHƯƠNG PHÁP TRÍCH XUẤT ĐẶC TRƯNG THỊ GIÁC VÀ TĂNG CƯỜNG DỮ LIỆU CHUYỂN ĐỘNG MẮT

Chương 3 trình bày các giải pháp đề xuất nhằm thứ nhất là chuẩn bị dữ liệu và thứ hai là tăng cường dữ liệu, chuẩn bị nguồn dữ liệu đầu vào cho các mô hình phân tích. Nội dung chương được cấu trúc dựa trên hai nhóm phương pháp có tính bổ trợ: (1) Phương pháp trích xuất tự động đặc trưng hành vi thị giác dựa trên chuỗi tọa độ điểm nhìn đối với các kích thích thị giác từ chuỗi khung hình động, nhằm định lượng quá trình nhận thức; và (2) Phương pháp tăng cường dữ liệu quỹ đạo ánh nhìn dựa trên thuật toán tối ưu tiến hóa, nhằm khắc phục giới hạn về quy mô của dữ liệu lâm sàng và cải thiện khả năng khái quát của các mô hình học máy.

Tập hợp các giải pháp này thực hiện chuyển đổi dữ liệu thô thành các véc tơ đặc trưng có cấu trúc. Bước xử lý này giúp tạo dữ liệu để triển khai các mô hình phân loại và hệ thống hỗ trợ can thiệp ở những chương tiếp theo.

3.1 Giới thiệu

3.1.1 Vai trò và rào cản trong phân tích dữ liệu chuyển động mắt

Dữ liệu chuyển động mắt thô chưa cung cấp trực tiếp các ý nghĩa lâm sàng. Do đó, quá trình trích xuất đặc trưng thị giác, bản chất là việc lượng hóa chuỗi tọa độ thành các tham số hành vi, đóng vai trò cơ sở trong các nghiên cứu về trẻ rối loạn phát triển [51]. Cụ thể, đối với nhóm trẻ rối loạn phổ tự kỉ, quá trình này hỗ trợ phát hiện các mẫu hành vi né tránh hoặc sự suy giảm chú ý vào các tín hiệu xã hội. Đối với nhóm trẻ rối loạn tính toán, quá trình này giúp xác định những thời điểm gia tăng tải nhận thức thông qua việc phân tích ma trận chuyển đổi ánh nhìn. Đối với nhóm trẻ mắc chứng khó đọc, quá trình này cung cấp các chỉ số về tần suất nhìn ngược để phân biệt rào cản nhận diện từ vựng với năng lực đọc hiểu [32].

Mặc dù quy trình tiền xử lý tại Chương 2 đã xây dựng được bộ dữ liệu thô đồng nhất, việc khai thác thông tin từ nguồn dữ liệu này vẫn đối mặt với hai rào cản kỹ thuật chính:

- Giới hạn trong việc trích xuất đặc trưng không gian động: Hiện nay, các nghiên cứu thường xác định vùng quan tâm thông qua phương pháp gán nhãn thủ công hoặc thiết lập ranh giới không gian tĩnh. Khi áp dụng vào chuỗi khung hình động, các phương pháp này bộc lộ hạn chế do đối tượng liên tục di chuyển, thay đổi kích thước hoặc bị che khuất, từ đó dẫn đến các sai số đáng kể [47].
- Giới hạn về quy mô và sự phân bố dữ liệu: Quá trình thu nhận dữ liệu hành vi thị giác từ nhóm trẻ rối loạn phát triển gặp nhiều khó khăn thực tế, dẫn đến các bộ dữ liệu có kích thước nhỏ và thiếu cân bằng. Bên cạnh đó, các phương pháp tổng hợp dữ liệu truyền thống thường chỉ đạt mức độ tương đồng ở không gian vĩ mô và bộc lộ hạn chế tại cấp độ vi chuyển động. Hệ quả là chuỗi tọa độ sinh ra không thỏa mãn ngưỡng dung sai không gian tự nhiên của nhãn cầu.

3.1.2 Động lực và đóng góp

Nhằm khắc phục các rào cản kỹ thuật đã nêu, luận án đề xuất một quy trình xử lý dữ liệu có tính hệ thống, bao gồm các đóng góp chính sau:

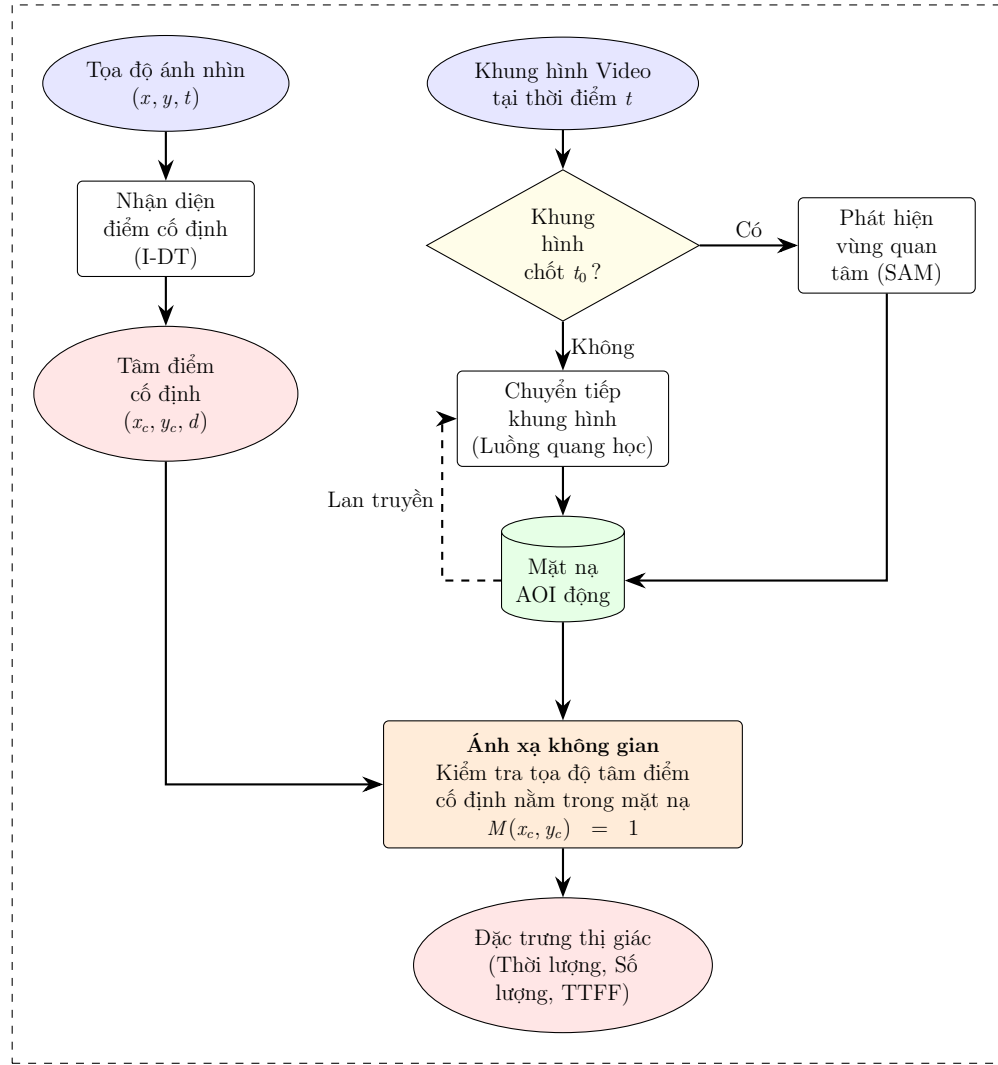
- Đề xuất phương pháp trích xuất đặc trưng tự động: Phương pháp này tích hợp thuật toán nhận diện điểm cố định (Python implementation of eye movements metrics & visualizations - PeyeMMV) [125], mô hình phân đoạn bất kì (Segment Anything Model - SAM) [126] và kỹ thuật luồng quang học (Optical Flow) [127]. Cơ chế kết hợp này hỗ trợ tính toán và cập nhật liên tục ranh giới của các vùng quan tâm động, qua đó khắc phục hạn chế của quy trình gán nhãn thủ công.
- Đề xuất phương pháp gia dữ liệu chuyển động mắt: Bằng cách kết hợp mô hình chuyển động sinh học và thuật toán tối ưu hóa. Cụ thể, phương pháp sử dụng mô hình chuyển động Brownian phân số (fBm) để mô phỏng các cử động mắt tự nhiên của con người. Dữ liệu này sau đó được tinh chỉnh và kiểm soát thông qua thuật toán di truyền (GA), kết hợp cùng thuật toán PeyeMMV đóng vai trò là bộ tiêu chuẩn đánh giá. Sự kết hợp này tạo ra một quy trình để tự động tạo ra các bộ dữ liệu tăng cường duy trì độ chân thực và mức độ tin cậy.

3.2 Phương pháp lai trích xuất đặc trưng thị giác tự động

Việc tính toán các tham số hành vi thị giác là bước xử lý nền tảng trong công nghệ theo dõi chuyển động mắt. Để xác định điểm cố định của ánh nhìn, hệ thống thường sử dụng hai nhóm thuật toán là phương pháp dựa trên ngưỡng vận tốc (Identification by Velocity Threshold - I-VT) [54] và phương pháp dựa trên ngưỡng phân tán (Identification by Dispersion Threshold - I-DT) [54]. Kỹ thuật I-VT dễ gây ra sai số do các cử động đầu tự phát của trẻ rối loạn phát triển làm thay đổi vận tốc góc. Ngược lại, thuật toán I-DT tính toán mức độ hội tụ không gian dựa trên công thức 3.1 giúp duy trì sự ổn định khi xuất hiện các tín hiệu nhiễu.

$$D = [\max(x) - \min(x)] + [\max(y) - \min(y)] \quad (3.1)$$

Dựa trên cơ sở đó, phương pháp trích xuất đặc trưng tự động đề xuất sự kết hợp giữa thuật toán PeyeMMV (thuộc nhóm I-DT) và SAM. Cấu trúc phương pháp được mô tả tại sơ đồ Hình 3.1.



Hình 3.1: Sơ đồ luồng phương pháp trích xuất đặc trưng thị giác tự động sử dụng kiến trúc lai (SAM + Thuật toán luồng quang học) và thuật toán PeyeMMV.

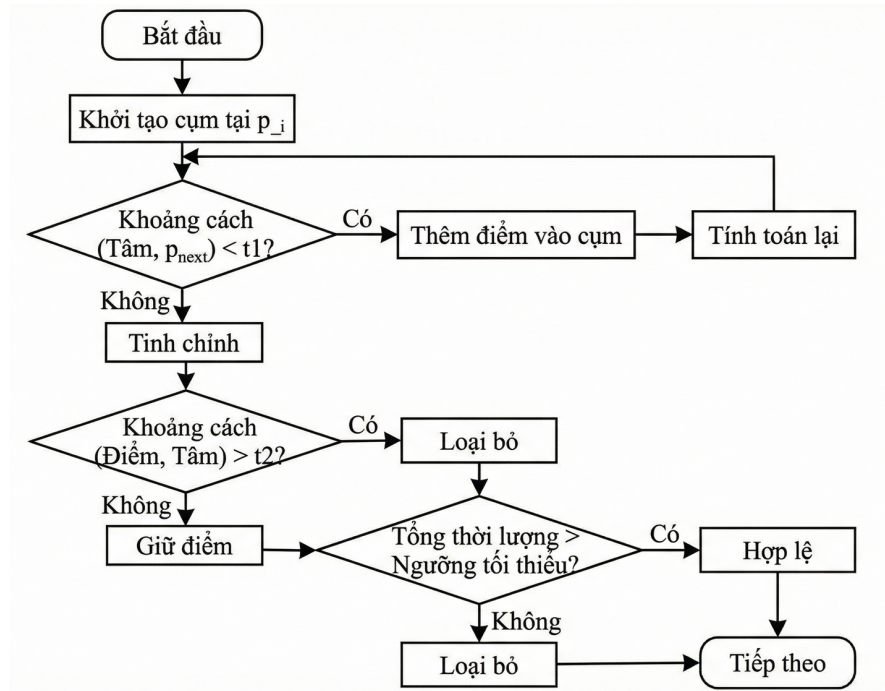
3.2.1 Bước 1: Phát hiện điểm cố định bằng thuật toán PeyeMMV

Luận án sử dụng thuật toán PeyeMMV, được tùy biến từ nền tảng mã nguồn mở eyeMMV [125], thuộc nhóm thuật toán dựa trên độ phân tán. Thuật toán này được lựa chọn nhằm phù hợp với thông số thiết bị thu thập và đặc điểm của trẻ RLPT thông qua hai yếu tố:

- Khả năng hoạt động ổn định ở tần số lấy mẫu thấp: Các thuật toán phân tích dựa trên tốc độ di chuyển của mắt thường yêu cầu thiết bị phần cứng phải đạt tần số thu thập dữ liệu lớn (> 250 Hz). Khi hệ thống sử dụng dòng thiết bị Tobii hoạt động ở mức 90 Hz, việc đo

lượng sự thay đổi tốc độ ánh nhìn thường phát sinh nhiều tín hiệu nhiễu. PeyeMMV khắc phục hạn chế này bằng cách tính toán cụm phân tán không gian trong cửa sổ thời gian trượt, giúp duy trì sự ổn định mà không yêu cầu tần số lấy mẫu cao.

- Khả năng lọc nhiễu do cử động: Trẻ RLPT thường có những cử động vùng đầu không ổn định trong quá trình thực hiện nhiệm vụ. Để giải quyết vấn đề này, thuật toán PeyeMMV với cơ chế lọc ngưỡng kép, chấp nhận các sai số vị trí nhỏ do rung lắc, giúp chuỗi tọa độ không bị ngắt quãng.



Hình 3.2: Sơ đồ mô tả quy trình 3 bước của thuật toán PeyeMMV

Thuật toán PeyeMMV (minh họa tại Hình 3.2) được triển khai dựa trên các ngưỡng không gian (t_1 , t_2) kết hợp với ngưỡng thời gian tối thiểu. Dữ liệu đầu vào là danh sách các tọa độ điểm nhìn (x, y, t). Quy trình thực hiện qua ba bước:

- Bước 1: Phân cụm sơ bộ: Thuật toán khởi tạo cụm điểm cố định từ điểm nhìn đầu tiên và liên tục cập nhật tọa độ trung tâm của cụm dữ liệu này. Khi hệ thống nhận được một tọa độ ánh nhìn tiếp theo, khoảng cách không gian từ điểm nhìn mới này đến vị trí trung tâm của cụm hiện tại được đo lường. Nếu khoảng cách đó nhỏ hơn một ngưỡng phân tán đã được cài đặt trước (t_1), điểm nhìn mới sẽ được tích hợp vào cụm.

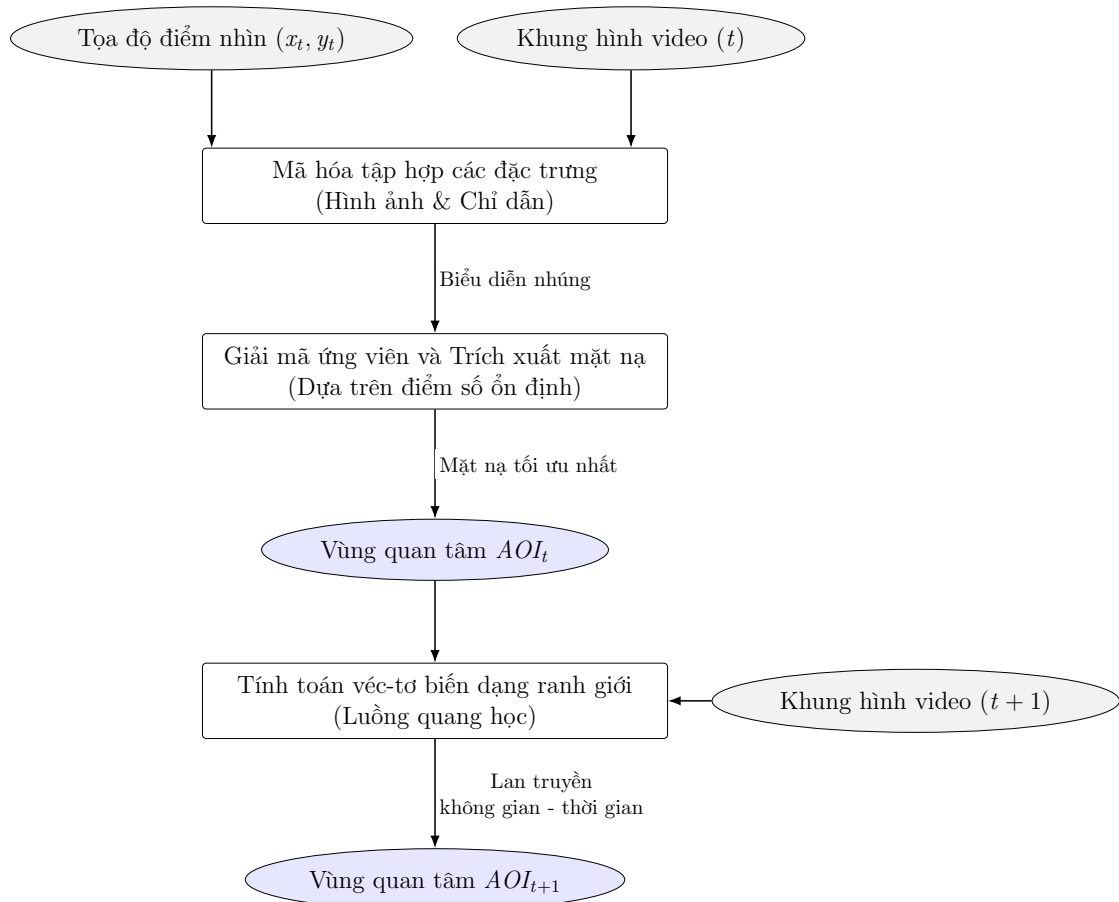
- Bước 2: Lọc tín hiệu nhiễu và tinh chỉnh cụm dữ liệu: Thuật toán tiến hành rà soát và tự động lọc các điểm nhìn có khoảng cách không gian tính đến tọa độ trung tâm vượt quá một giới hạn lọc nhiễu t_2 đã được thiết lập. Sau quá trình sàng lọc các điểm dữ liệu bất thường này, vị trí trung tâm mới của cụm (x_c, y_c) và thời lượng duy trì ảnh nhìn tại vùng đó được tính toán lại.
- Bước 3: Xác thực ngưỡng thời gian: Mỗi cụm được đánh giá dựa trên tổng thời gian tồn tại của mỗi cụm d . Nếu d thấp hơn ngưỡng, cụm được phân loại là chuyển động đường trượt và bị loại bỏ. Các cụm đạt điều kiện được trích xuất làm đầu vào cho khối hợp nhất dữ liệu.

3.2.2 Bước 2: Xác định và theo dõi vùng quan tâm động

3.2.2.1 Cơ chế phân đoạn và lan truyền không gian và thời gian

Để nhận diện các vùng quan tâm trên luồng dữ liệu động, luận án đề xuất một chuỗi quy trình kết hợp giữa khả năng phân đoạn không gian của mô hình SAM [126] và khả năng theo dõi chuỗi thời gian của thuật toán luồng quang học [127]. Tọa độ điểm nhìn của đối tượng được hệ thống sử dụng trực tiếp làm tín hiệu dẫn hướng để trích xuất vật thể [128]. Quy trình tổng thể được minh họa trong Hình 3.3 và bao gồm các giai đoạn như sau:

- Khởi tạo tập hợp các đặc trưng: Hệ thống đồng thời tiếp nhận hình ảnh trực quan từ khung hình video và tọa độ điểm nhìn (x_t, y_t) của đối tượng tại thời điểm t . Khối mã hóa của mô hình SAM trích xuất ma trận đặc trưng hình ảnh và nhúng tọa độ không gian thành các véc tơ chỉ dẫn. Cơ chế này hỗ trợ hệ thống phân tích cấu trúc bối cảnh xung quanh điểm tập trung thị giác.
- Giải mã và trích xuất vùng quan tâm: Khối giải mã tiếp nhận dữ liệu nhúng để tạo ra các mặt nạ ứng viên bao hàm nhiều mức độ chi tiết (từ một thành phần nhỏ đến toàn bộ cấu trúc vật thể). Nhằm xác định vùng quan tâm, hệ thống thực thi cơ chế đánh giá nội bộ để tính toán điểm số ổn định của từng mặt nạ. Khu vực đạt độ bền vững cấu trúc tốt nhất được chọn lọc và xuất ra thành vùng quan tâm AOI_t .
- Lan truyền vùng quan tâm theo chuỗi thời gian: Nhằm tối ưu hóa tài nguyên tính toán thay vì thực thi lại mô hình phân đoạn trên từng khung hình riêng lẻ, hệ thống áp dụng thuật toán luồng quang học để



Hình 3.3: Quy trình phân đoạn và lan truyền vùng quan tâm (AOI) theo chuỗi thời gian

duy trì tính liên tục của dữ liệu. Thuật toán tiếp nhận kết quả vùng quan tâm AOI_t đã được xác lập cùng dữ liệu hình ảnh tại hai thời điểm liên tiếp t và $t + 1$. Bằng cách phân tích trường chuyển động cục bộ của các điểm ảnh giữa hai khung hình, thuật toán thiết lập một hệ véc tơ biến dạng ranh giới. Căn cứ vào hệ véc tơ này, thuật toán tiến hành tịnh tiến và điều chỉnh hình thái của AOI_t để kết xuất thành vùng quan tâm AOI_{t+1} cho thời điểm hiện tại. Quá trình chuyển tiếp này hỗ trợ hệ thống theo dõi mục tiêu xuyên suốt khi đối tượng thay đổi hướng nhìn hoặc vật thể dịch chuyển vị trí trên giao diện.

3.2.2.2 Thuật toán tổng thể phương pháp lai trích xuất đặc trưng

Thuật toán 3.1 Quy trình lai trích xuất đặc trưng thị giác (SAM + Optical Flow + PeyeMMV)

Input: Chuỗi điểm nhìn thô $G = \{(x_t, y_t, \tau_t)\}$, Video bối cảnh $V = \{v_t\}$

Output: Véc tơ đặc trưng thị giác ngữ nghĩa $F_{semantic}$

```

/* 1. Lọc nhiễu và nhận diện điểm cố định (PeyeMMV) */
1  $Fixations \leftarrow$  I-DT_Algorithm( $G$ , Thresholddispersion, Thresholdtime)
/* 2. Phân đoạn và lan truyền vùng quan tâm (AOI) */
2  $AOI\_List \leftarrow \emptyset$  foreach khung hình  $v_t \in V$  do
3   if  $t$  là khung hình chính (Keyframe) then
4      $Mask_t \leftarrow$  SAM_Encoder_Decoder( $v_t$ ,  $Fixations(t)$ )
5   else
6      $Mask_t \leftarrow$  OpticalFlow_Update( $Mask_{t-1}$ ,  $v_{t-1}$ ,  $v_t$ )
7   end if
8    $AOI\_List.append(Mask_t)$ 
9 end foreach
/* 3. Ánh xạ không gian và trích xuất đặc trưng */
10  $F_{semantic} \leftarrow \emptyset$  foreach tâm điểm cố định  $(x_c, y_c) \in Fixations$  do
11   if  $Mask_t(x_c, y_c) == 1$  then
12      $F_{semantic} \leftarrow F_{semantic} \cup \{ExtractMetrics(x_c, y_c, Mask_t)\}$ 
13   end if
14 end foreach
15 return  $F_{semantic}$ 

```

Thuật toán 3.1 trình bày cấu trúc tổng thể của quy trình tính toán thuộc phương pháp lai đề xuất. Tiến trình này biểu thị sự tương tác trực tiếp giữa hệ thống tọa độ dữ liệu nguyên bản và cơ chế phân đoạn kết hợp không gian cùng thời gian.

3.2.3 Bước 3: Ánh xạ ngữ nghĩa và trích xuất đặc trưng chiến lược

Nhằm phục vụ mục tiêu phân tích hành vi, quá trình trích xuất dữ liệu được hệ thống phân tách thành hai cấp độ đặc trưng:

- Nhóm đặc trưng cơ sở: Bao gồm các tham số được trích xuất trực tiếp từ dữ liệu thô như thời lượng điểm cố định, số lượng điểm cố định và thời gian đến điểm cố định đầu tiên. Nhóm tham số này chỉ đo lường đặc tính sinh lý của mắt tại các vị trí tọa độ hình học, do đó chưa phản ánh được tiến trình nhận thức vì thiếu thông tin về ý nghĩa của khu vực được quan sát.
- Nhóm đặc trưng chiến lược: Bao gồm các tham số được thiết lập thông qua phép ánh xạ tọa độ điểm nhìn vào các vùng quan tâm động. Quá trình xử lý này gắn kết dữ liệu hành vi thị giác với ý nghĩa của từng vùng nội dung, từ đó cung cấp các chỉ số định lượng để đánh giá cách thức đối tượng tiếp nhận và xử lý thông tin.

Dựa trên đặc điểm của từng nhóm rối loạn, ba bộ đặc trưng mang tính phân biệt được đề xuất như sau:

1. Entropy quỹ đạo quét của trẻ RLPTK:

Đây là chỉ số phân tán quỹ đạo nhìn cho trẻ RLPTK. Để đo lường sự thay đổi trong cách trẻ tìm kiếm thông tin khi quan sát, hệ thống tính toán chỉ số phân tán không gian. Gọi p_i là xác suất điểm nhìn rơi vào vùng AOI_i , mức độ hỗn loạn được tính theo công thức Shannon [129]:

$$H = - \sum_{i=1}^k p_i \log_2(p_i) \quad (3.2)$$

Các quỹ đạo có tính định hướng và mục tiêu rõ ràng tạo ra giá trị H thấp, trong khi các quỹ đạo phân tán, thiếu định hướng (đặc trưng ở trẻ RLPTK) tạo ra giá trị H cao hơn. Đặc trưng này cung cấp dữ liệu để các mô hình học máy phát hiện và phân loại hành vi quan sát của trẻ.

2. Ma trận chuyển đổi vùng quan tâm của trẻ rối loạn tính toán:

Trẻ RLTT thường gặp trở ngại trong việc kết nối các thành phần thông tin của bài toán. Hệ thống trích xuất *ma trận chuyển đổi* $M_T \in \mathbb{R}^{n \times n}$, trong đó phần tử $m_{i,j}$ biểu diễn tần suất ánh nhìn di chuyển từ AOI_i sang AOI_j . Ma trận này có thể coi là một bảng thống kê ghi nhận tần suất ánh nhìn di chuyển qua lại giữa các vùng nội dung khác nhau. Đặc trưng này hỗ trợ phát hiện chiến lược giải bài của trẻ như đọc lướt hay liên tục đối chiếu thông tin, qua đó làm cơ sở để phân tích tình trạng quá tải nhận thức.

3. Tỷ lệ bước nhảy mắt hồi quy của trẻ mắc chứng khó đọc:

Nhóm trẻ mắc chứng khó đọc thường gặp trở ngại trong giải mã âm vị. Trong quá trình đọc văn bản, chuyển động mắt tiến thể hiện mức độ trôi chảy, trong khi chuyển động mắt hồi quy báo hiệu trẻ đang gặp khó khăn trong xử lý nhận thức. Hệ thống trích xuất tỉ lệ hồi quy [32] (R_{reg}) như sau:

$$R_{reg} = \frac{N_{reg}}{N_{prog} + N_{reg}} \quad (3.3)$$

Trong đó N_{prog} và N_{reg} lần lượt là số lượng bước nhảy tiến và lùi. Đặc trưng động học này cung cấp thông tin đầu vào cho các mô hình phân tích hành vi đọc, giúp phân biệt giữa rào cản nhận diện mặt chữ và các yếu tố ảnh hưởng khác.

3.3 Thực nghiệm và đánh giá phương pháp trích xuất đặc trưng thị giác tự động

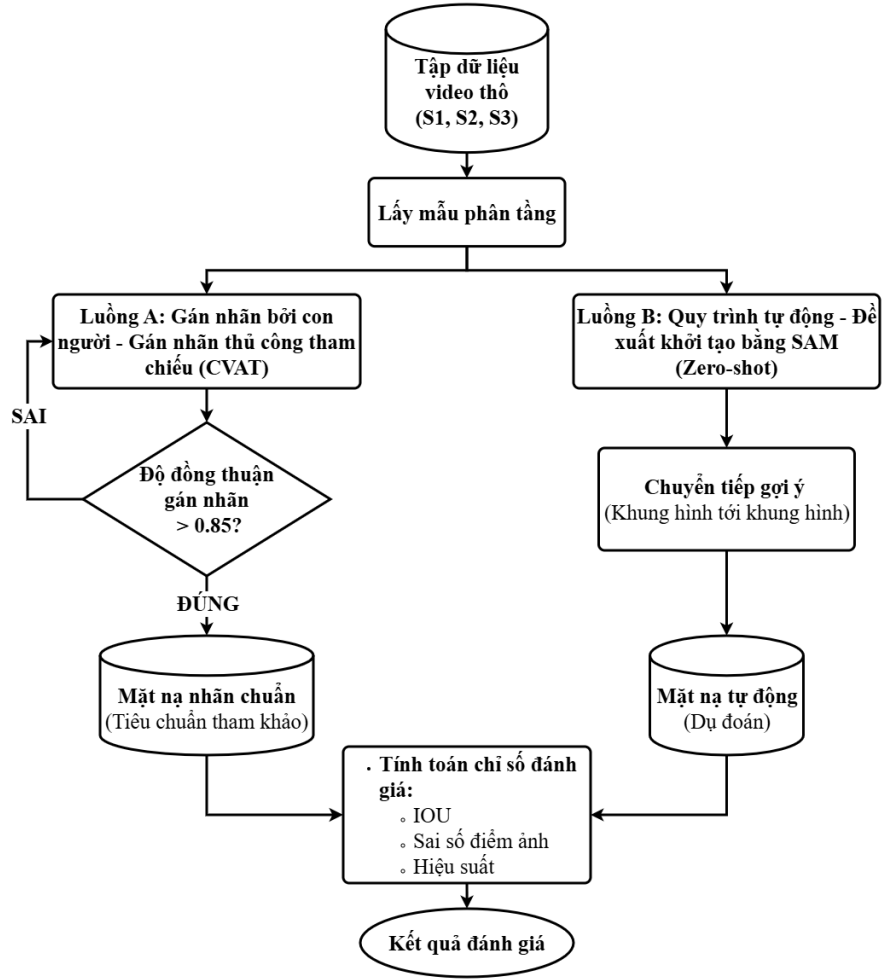
Kế thừa nguồn dữ liệu thô đã được thiết lập thông qua quy trình thu thập tại mục 2.3 và 2.4, nội dung phần này tập trung trình bày việc thử nghiệm phương pháp trích xuất đặc trưng thị giác tự động. Mục tiêu trọng tâm là mô tả chi tiết các bước xử lý và chuyển đổi dữ liệu tọa độ điểm nhìn thành các chỉ số hành vi định lượng, qua đó cung cấp dữ liệu đầu vào cho các mô hình phân tích ở những giai đoạn tiếp theo.

3.3.1 Đánh giá kỹ thuật nhận diện vùng quan tâm tự động

3.3.1.1 Hiệu năng phân đoạn vùng quan tâm

Để kiểm tra hiệu quả của phương pháp phân đoạn, luận án thiết lập quy trình kiểm chứng chéo (Hình 3.4) bằng cách so sánh trực tiếp kết quả bộ dữ liệu của phương pháp đề xuất với bộ dữ liệu đối chứng do các chuyên gia tự khoanh vùng thủ công. Mức độ chính xác được đo lường dựa trên tỷ lệ trùng khớp giữa hai kết quả này. Quy trình kiểm tra được chia thành hai luồng xử lý độc lập. Cụ thể, quy trình bao gồm hai luồng như sau:

- Luồng A: Gán nhãn thủ công thông qua CVAT [130] để tạo dữ liệu đối chứng, với điều kiện các mẫu phải đạt hệ số đồng thuận giữa các chuyên gia Cohen's $\kappa > 0,85$.



Hình 3.4: Quy trình xác thực khả năng nhận diện vùng quan tâm tự động

– Luồng B: Thực hiện phân đoạn tự động bằng SAM.

Hiệu suất của thuật toán được lượng hóa thông qua chỉ số Jaccard (Intersection over Union - IoU) và hệ số Dice [131]:

$$IoU = \frac{|A \cap B|}{|A \cup B|}, \quad Dice = \frac{2|A \cap B|}{|A| + |B|} \quad (3.4)$$

Thử nghiệm thực tế trên bộ dữ liệu gồm 1480 khung hình cho thấy mô hình đạt chỉ số IoU trung bình $0,8703 \pm 0,1086$ và Dice Score $0,9268 \pm 0,0675$. Sai số định vị điểm ảnh trung bình đạt mức $5,05 \pm 4,99$ pixel. Các số liệu này đều vượt qua ngưỡng tiêu chuẩn tối thiểu là 80% thường được yêu cầu đối với các hệ thống phân tích dữ liệu y khoa [131]. Kết quả định lượng cho thấy phương pháp phân đoạn tự động theo quy trình đề xuất duy trì được

độ trung thực không gian, đáp ứng yêu cầu kỹ thuật để thay thế quy trình gán nhãn thủ công trong các giai đoạn phân tích tiếp theo.

3.3.1.2 Cơ chế hợp nhất dữ liệu và tối ưu hóa tính toán

Sau khi hoàn tất quá trình trích xuất, dữ liệu được chuyển đến khối *ánh xạ không gian*. Hệ thống thực hiện kiểm tra giao cắt hình học giữa tọa độ điểm cố định và mặt nạ nhị phân để xác định vùng chú ý. Tại đây, hệ thống kiểm tra xem tọa độ điểm nhìn có nằm bên trong vùng mặt nạ đối tượng hay không. Việc sử dụng mặt nạ bám sát từng điểm ảnh thay vì dùng các khung viền chữ nhật giúp giảm thiểu sai số khi ánh nhìn rơi vào vùng biên của đối tượng. Điều kiện ánh xạ hợp lệ [132] được xác định khi:

$$M(x_c, y_c) = 1 \quad (3.5)$$

Để đáp ứng yêu cầu hoạt động thời gian thực, khối lượng tính toán của hệ thống cần được tối ưu hóa. Nếu xử lý mô hình phân đoạn hình ảnh trên mọi khung hình liên tiếp, với N là số lượng token và D là chiều của véc tơ nhúng, cơ chế tự chú ý sẽ dẫn đến độ phức tạp tính toán tăng theo hàm bậc hai [133], hệ thống khó duy trì được tốc độ tiêu chuẩn:

$$C_{vanilla} = \mathcal{O}(K \cdot N^2 \cdot D) \quad (3.6)$$

Khối lượng tính toán này vượt quá giới hạn tài nguyên phần cứng, khiến hệ thống khó đạt được tốc độ xử lý tiêu chuẩn. Để khắc phục hạn chế trên, một kỹ thuật xử lý luân phiên được áp dụng trong hệ thống. Cụ thể, gọi K là chu kỳ nội suy, SAM chỉ được kích hoạt tại khung hình chính t_0 . Tại $K - 1$ khung hình trung gian tiếp theo, hệ thống chuyển sang vận hành thuật toán *luồng quang học* dựa trên đạo hàm cường độ sáng cục bộ để theo dõi và cập nhật vị trí đối tượng. Giải pháp này giúp giảm tải cho bộ nhớ đồ họa và đảm bảo độ trễ thấp cho toàn hệ thống. Chi phí tính toán lúc này được tối ưu hóa thành [134]:

$$C_{hybrid} = \mathcal{O}(N^2 \cdot D) + \mathcal{O}((K - 1) \cdot N) \quad (3.7)$$

Việc giảm độ phức tạp từ bậc hai xuống tuyến tính giúp giảm tải bộ nhớ GPU, đưa độ trễ hệ thống về mức phù hợp cho xử lý thời gian thực.

3.3.1.3 Đánh giá độ trễ và khả năng đáp ứng thời gian thực

Hệ thống được thử nghiệm trên cấu hình máy tính gồm vi xử lý Intel Core i7 và card đồ họa NVIDIA RTX 5060. Kết quả thực nghiệm qua các giai đoạn cấu hình được ghi nhận như sau:

- Khi vận hành mô hình SAM độc lập, thời gian xử lý trung bình của hệ thống là 157,18 ms/khung hình (tương đương tốc độ 6,4 FPS).
- Khi áp dụng kiến trúc lan truyền lai ($K = 10$), thời gian xử lý giảm xuống còn 30,27 ms/khung hình (tương đương tốc độ 33 FPS).
- Khi tích hợp thêm thuật toán PeyeMMV tốc độ xử lý tổng thể của hệ thống duy trì ở mức xấp xỉ 31,5 FPS.

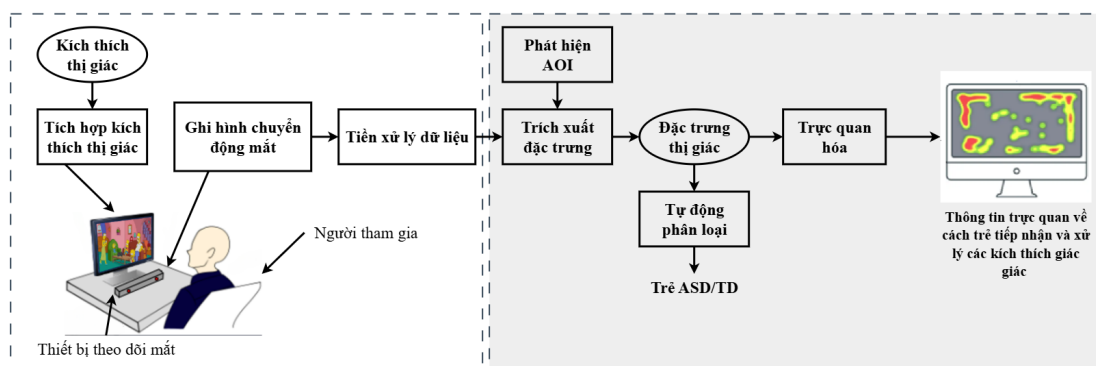
Tốc độ tổng thể đạt trên 30 FPS, đáp ứng tiêu chuẩn vận hành thời gian thực cho hệ thống vòng kín [135], đủ biên độ an toàn tương đối hẹp (khoảng 5%). So với phương pháp gán nhãn thủ công với thời gian trung bình 45 giây/khung hình, phương pháp tự động giúp giảm thiểu xử lý dữ liệu. Đối với video có vật thể di chuyển, phương pháp mặt nạ động vừa duy trì điều kiện thời gian thực, vừa nâng cao độ chính xác phân đoạn từ 65% lên 88% so với phương pháp AOI tĩnh [136].

3.3.2 Thiết lập thực nghiệm phân tích và phân loại lâm sàng

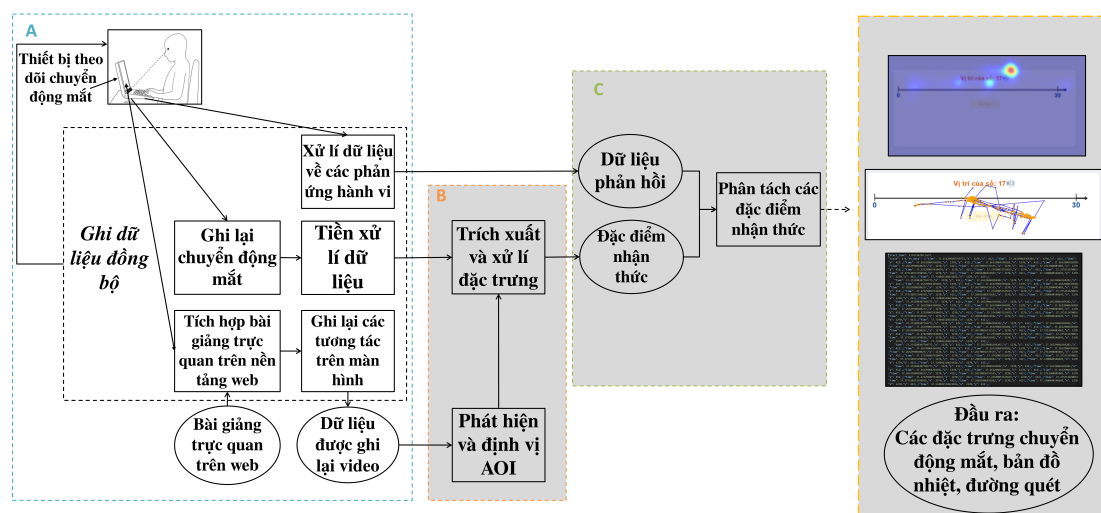
Sau khi các đặc trưng thị giác cơ bản bao gồm thời lượng cố định mắt, số lượng điểm cố định tại vùng quan tâm, và thời gian trễ đến điểm cố định đầu tiên (Time to First Fixation - TTFF), được trích xuất, hệ thống phân luồng xử lý dữ liệu tùy theo từng nhóm đối tượng.

Trong hệ thống thu thập dữ liệu và phân tích đặc trưng thị giác của trẻ RLPTK như minh họa tại Hình 3.5, các đặc trưng thị giác đóng vai trò là đầu vào cho hai mô-đun chính: trực quan hóa dữ liệu và phân loại tự động thông qua mô hình học máy.

Đối với hệ thống thu thập dữ liệu và phân tích đặc trưng nhận thức của trẻ RLTT được minh họa tại Hình 3.6, các đặc trưng thị giác được đưa vào mô-đun phân tích đặc trưng nhận thức. Mô-đun này kết hợp với các nguồn dữ liệu tương tác để phân tích sâu hơn về hành vi thị giác và cơ chế chú ý của trẻ.



Hình 3.5: Hệ thống thu thập và trích xuất đặc trưng thị giác cho trẻ RLPTK



Hình 3.6: Hệ thống thu thập và trích xuất đặc trưng thị giác cho trẻ có rối loạn tính toán

Để đảm bảo tính khách quan và độ tin cậy của các kết quả thực nghiệm trên cỡ mẫu y sinh quy mô nhỏ và vừa, quy trình đánh giá thống kê được thực hiện qua ba kỹ thuật cơ bản:

1. *Kiểm chứng chéo theo cấp độ đối tượng*: Khác với dữ liệu văn bản hay hình ảnh thông thường, chuỗi dữ liệu sinh trắc học⁸ có tính liên kết thời gian chặt chẽ. Nếu chia ngẫu nhiên dữ liệu của cùng một trẻ vào cả tập huấn luyện và tập kiểm thử, mô hình sẽ mắc lỗi rò rỉ dữ liệu và ghi nhớ thói quen cá nhân thay vì nhận diện đúng đặc điểm rối loạn. Do đó, luận án áp dụng phương pháp kiểm chứng chéo luân phiên từng đối tượng (*Leave-One-Subject-Out Cross-Validation - LOSOCV*) [137]. Tại mỗi vòng lặp, toàn bộ dữ liệu của một cá thể được tách ra làm tập kiểm thử, mô hình được huấn luyện trên $N - 1$ cá thể còn lại. Quy trình này lặp lại N lần, đảm

bảo hiệu năng báo cáo là hiệu năng tổng quát hóa thực sự trên dữ liệu.

2. *Ước lượng khoảng tin cậy 95% (95% Confidence intervals - CI)*: Với cỡ mẫu hạn chế, một giá trị độ chính xác điểm không đủ để đại diện cho độ tin cậy của toàn bộ quần thể. Luận án tính toán và báo cáo khoảng tin cậy 95% (95% CI) cho các chỉ số hiệu năng (Accuracy, Diện tích dưới đường cong ROC (Area Under the Curve - AUC)) bằng *phương pháp tái chọn mẫu Bootstrap* [138] với 1.000 lần lặp. Quy trình này xác lập biên độ sai số khoa học, chứng minh kết quả không bị sai lệch bởi các giá trị ngoại lai.

3. *Kiểm định ý nghĩa thống kê phi tham số*:

Với đặc điểm của dữ liệu sinh trắc học trên cỡ mẫu nhỏ, các biến số thường có xu hướng phân phối lệch và chưa đáp ứng giả định phân phối chuẩn. Do đó, luận án áp dụng các phương pháp kiểm định thống kê phi tham số trong quá trình phân tích dữ liệu [139]:

- Kiểm định Mann-Whitney U: Được ứng dụng nhằm đối sánh sự khác biệt về các đặc trưng định lượng giữa nhóm trẻ rối loạn phát triển và nhóm đối chứng [140]. Phép kiểm định này hỗ trợ đánh giá khả năng phân tách của các tham số thị giác thu thập từ hai nhóm phân tích.
- Hệ số tương quan hạng Spearman: Được ứng dụng nhằm lượng hóa mối liên hệ thống kê giữa các biến số trích xuất từ luồng dữ liệu đa phương thức [141]. Chỉ số này cung cấp cơ sở để phản ánh mức độ tương quan đồng thời giữa hành vi thị giác (thời lượng quan sát) và phản hồi tương tác (tốc độ thao tác chuột).
- Kiểm định *Tau – U*: Được ứng dụng nhằm phân tích chuỗi thời gian của mỗi trẻ, qua đó xác định xu hướng biến thiên hành vi của từng đối tượng xuyên suốt các pha thực nghiệm [142]. Phép kiểm định này hỗ trợ đánh giá mức độ không chồng lấp của dữ liệu thị giác khi trẻ chuyển đổi giữa các trạng thái kích thích khác nhau.

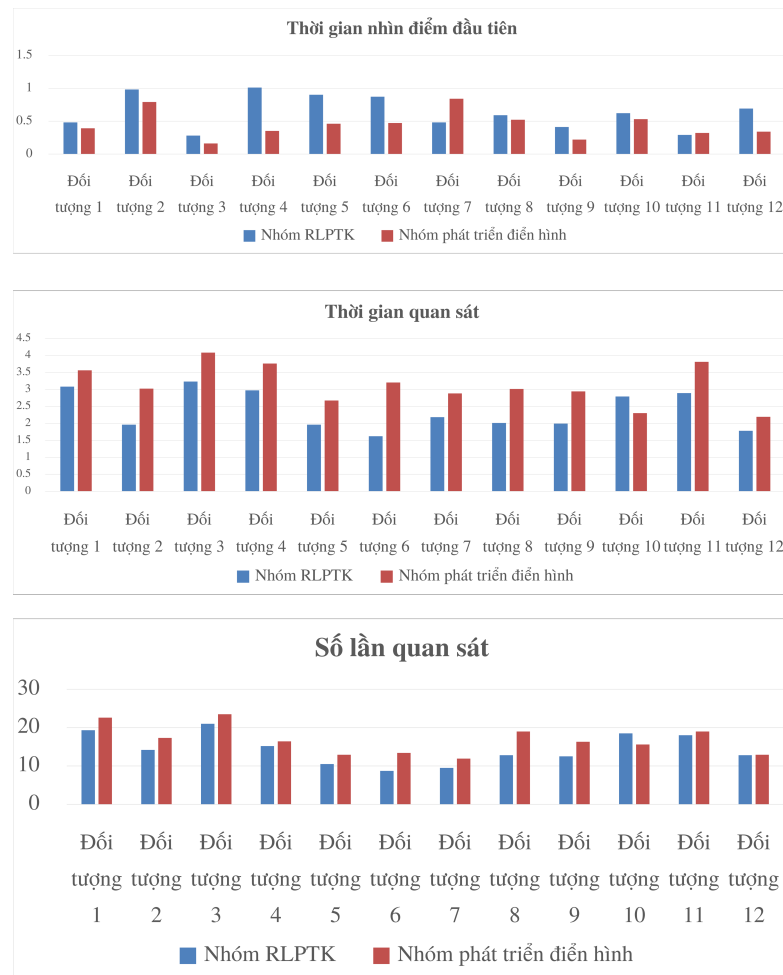
3.3.3 Đánh giá đặc trưng thị giác và hiệu năng hỗ trợ chẩn đoán ở trẻ rối loạn phổ tự kỉ

Quy trình đánh giá kết quả thực nghiệm đối với nhóm trẻ rối loạn phổ tự kỉ được thực hiện nhằm kiểm chứng độ tin cậy của phương pháp trích xuất

8. *Dữ liệu sinh trắc học*: Trong phạm vi luận án này, *dữ liệu sinh trắc học* được hiểu là các dấu ấn sinh học hành vi thu nhận từ quá trình tương tác, trong đó *dữ liệu sinh trắc học thị giác* đóng vai trò là nguồn dữ liệu đầu vào chủ đạo phản ánh quá trình xử lý nhận thức của trẻ.

đặc trưng và hiệu năng phân loại của hệ thống. Nội dung đánh giá được tổ chức thành năm bước chính bao gồm: (1) Phân tích thống kê và trực quan hóa dữ liệu hành vi thị giác nhằm nhận diện các quy luật phân bố sự chú ý; (2) Thiết lập quy trình trích xuất đặc trưng và tối ưu hóa cấu hình cho mô hình học máy; (3) Kiểm chứng hiệu năng phân loại và định lượng mức độ đóng góp của các đặc trưng hành vi; (4) Phân tích thành phần nhằm xác thực khả năng loại bỏ nhiễu động học của phương pháp lai; và (5) So sánh hiệu năng của hệ thống với các công trình nghiên cứu hiện tại.

3.3.3.1 Phân tích thống kê và trực quan hóa hành vi thị giác

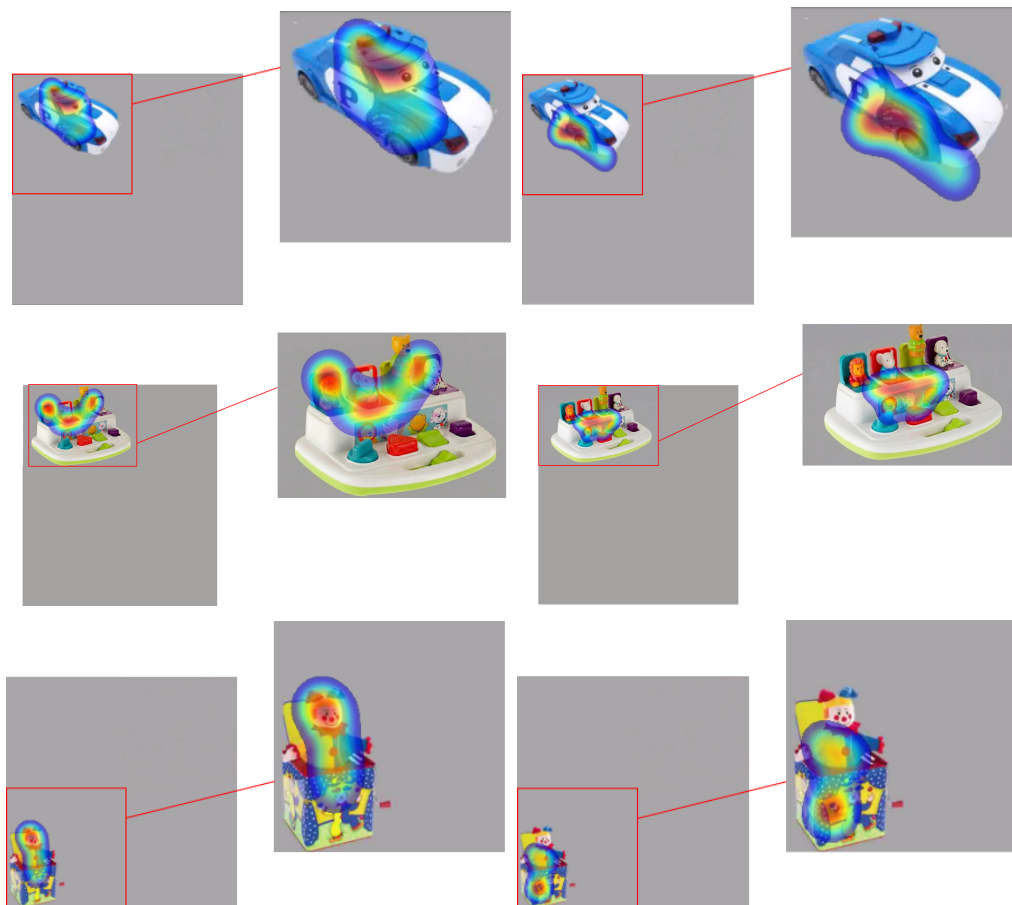


Hình 3.7: Kết quả các đặc trưng thống kê của chuyển động mắt

Các kết quả đặc trưng thống kê của chuyển động mắt được trình bày trên Hình 3.7 cho thấy nhóm trẻ RLPTK có số lượng điểm cố định thị giác thấp hơn so với trẻ PTĐH. Nhóm trẻ này cũng có xu hướng duy trì thời

gian quan sát ngắn hơn, phản ánh sự hạn chế trong việc tiếp nhận thông tin hình ảnh. Thêm vào đó, thời gian để trẻ bắt đầu tập trung ánh nhìn vào một điểm đầu tiên cũng kéo dài hơn.

Dữ liệu biểu diễn dưới dạng bản đồ nhiệt (Hình 3.8) chỉ ra sự khác biệt về cách phân bố ánh nhìn giữa hai nhóm. Trẻ RLPTK có xu hướng tập trung vào các thành phần phi xã hội (như nút bấm, thân hộp), trong khi hạn chế chú ý vào các vùng chứa khuôn mặt. Ngược lại, nhóm đối chứng ưu tiên chú ý vào các yếu tố giao tiếp xã hội.

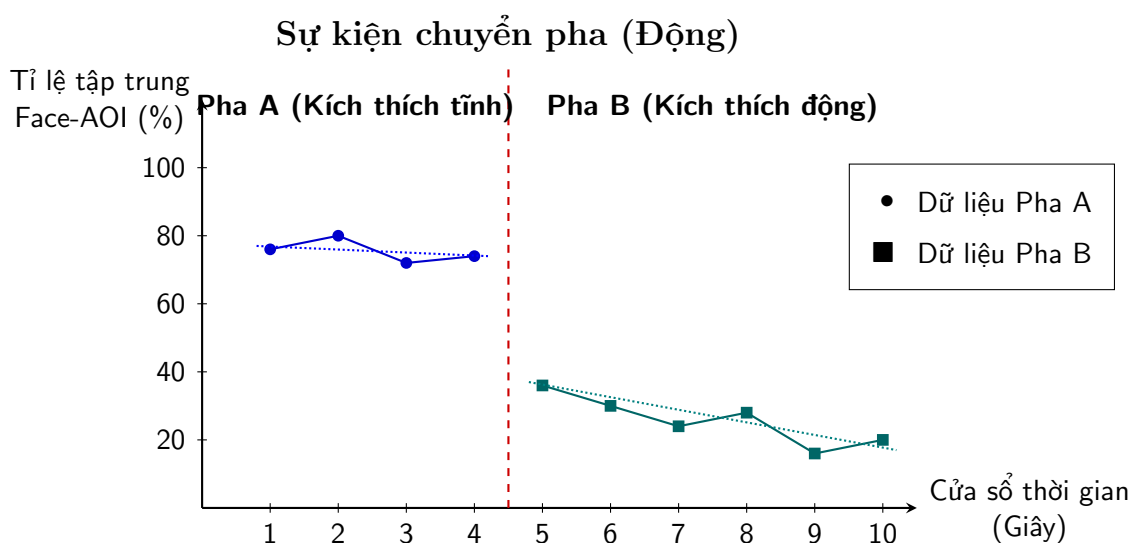


Hình 3.8: Bản đồ nhiệt của trẻ RLPTK (cột phải) và trẻ phát triển điển hình (cột trái) khi quan sát các hình ảnh có yếu tố xã hội

Để đánh giá độ tin cậy của các quan sát này, hệ thống thực hiện phép đo kiểm định thống kê độc lập. Kết quả xác nhận tỉ lệ thời gian cố định ánh nhìn tại vùng khuôn mặt ở nhóm RLPTK thấp hơn so với nhóm đối chứng ($p < 0,01$). Dữ liệu này phản ánh sự suy giảm chú ý vào các yếu tố xã hội, phù hợp với các quan sát từ bản đồ nhiệt.

Bên cạnh phương pháp phân tích tổng thể, luận án ứng dụng kiểm định $Tau - U$ để đánh giá chuỗi thời gian của từng trẻ nhằm theo dõi sự thay đổi mức độ tập trung khi hoàn cảnh giao tiếp thay đổi. Quá trình so sánh tỉ lệ duy trì điểm cố định tại vùng khuôn mặt được thực hiện giữa pha A (kích thích tĩnh) và pha B (kích thích động).

Kết quả kiểm định $Tau - U$ ghi nhận giá trị Tau dao động trong khoảng từ $-0,65$ đến $-0,82$ ($p < 0,05$). Các chỉ số âm cùng mức độ tương quan chặt chẽ này phản ánh một xu hướng thống kê đáng chú ý: sự xuất hiện của các kích thích động (Pha B) có mối liên hệ mật thiết với hiện tượng phân tán quỹ đạo chú ý tại những vùng tương tác xã hội. Cụ thể, khi các hình ảnh chuyển động được đưa vào màn hình, mức độ tập trung thị giác của nhóm trẻ rối loạn phát triển vào khu vực tương tác trọng tâm ghi nhận sự sụt giảm.



Hình 3.9: Biểu đồ minh họa chuỗi thời gian của một đối tượng.

Hình 3.9 minh họa phân tích chuỗi thời gian của một đối tượng điển hình. Tại pha A, dữ liệu duy trì mức ổn định từ 70% – 80%. Biểu đồ minh họa sự dịch chuyển tỷ lệ tập trung thị giác tại vùng khuôn mặt của một đối tượng trẻ rối loạn phát triển trước và sau khi xuất hiện kích thích động. Sự sụt giảm mức độ tập trung giữa hai pha tương ứng với giá trị $Tau - U$ âm mang ý nghĩa thống kê. Tại thời điểm chuyển pha, quỹ đạo thị giác thay đổi, giảm xuống dưới 40% và duy trì trạng thái phân tán trong pha B. Tỉ lệ không chồng lấp dữ liệu giữa hai pha đạt 100% (1, 0), tương ứng với giá trị $Tau - U = -0,82$. Kết quả thực nghiệm cho thấy hành vi thị giác của

nhóm trẻ rối loạn phát triển ghi nhận sự phân tán chú ý khi độ phức tạp của các kích thích xã hội gia tăng.

3.3.3.2 Thiết lập luồng trích xuất đặc trưng và tối ưu hóa mô hình phân loại

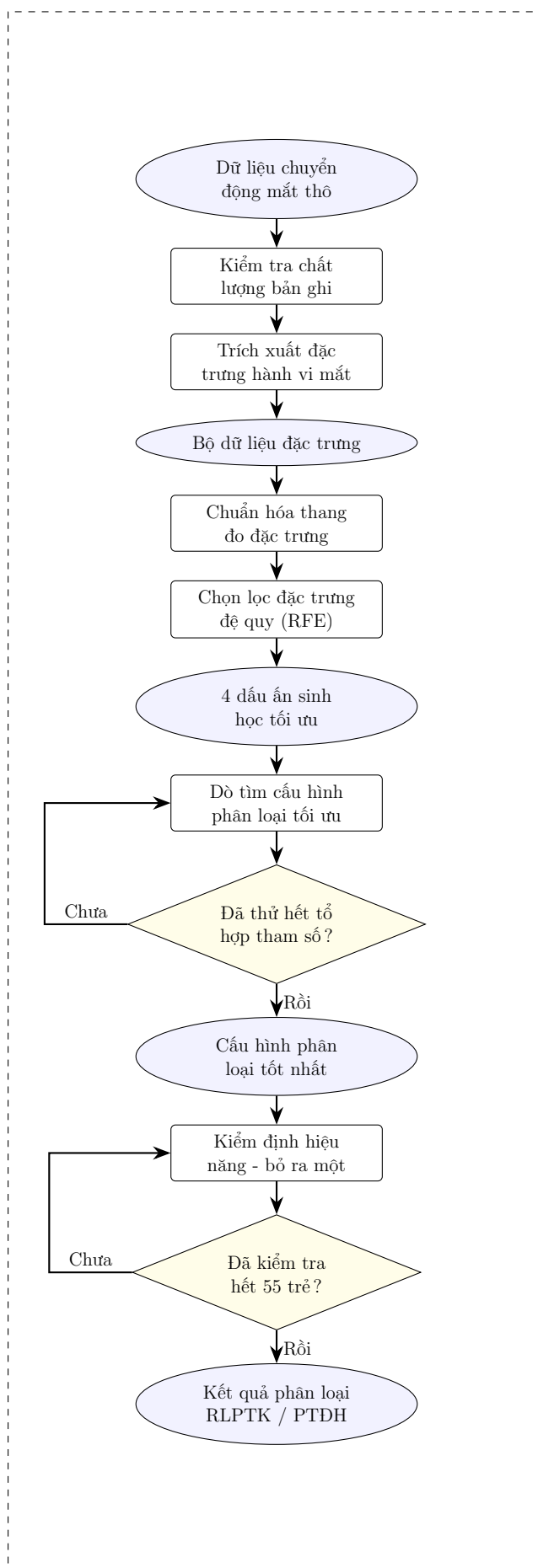
Để đảm bảo tính minh bạch và độ tin cậy của kết quả chẩn đoán, luận án xây dựng luồng xử lý tín hiệu thị giác và phân loại tự động (Hình 3.10). Quá trình này được thực hiện qua các bước như sau:

- Tiếp nhận và kiểm tra chất lượng dữ liệu: Hệ thống thu thập chuỗi tọa độ điểm nhìn từ 55 trẻ tham gia. Các bản ghi không đạt đủ số lượng điểm đo tối thiểu sẽ bị loại bỏ nhằm duy trì độ tin cậy cho nguồn dữ liệu đầu vào.
- Đồng bộ hóa dữ liệu cá nhân: Nhằm loại bỏ sự chênh lệch về tốc độ phản xạ tự nhiên giữa các cá nhân, hệ thống thực hiện hiệu chỉnh dữ liệu dựa trên giá trị cơ sở của từng trẻ. Tại pha định hướng với dấu thập trên màn hình, hệ thống đo tốc độ phản xạ cơ sở của mỗi trẻ để xác định giá trị trung bình $\mu_{base,i}$ và độ lệch chuẩn $\sigma_{base,i}$. Trong pha kích thích tiếp theo, chỉ số thời gian thô $X_{raw,i}$ được chuyển đổi sang điểm số tương đối Z (Intra subject Baseline Z scoring) theo công thức [143]:

$$Z_{feature,i} = \frac{X_{raw,i} - \mu_{base,i}}{\sigma_{base,i}} \quad (3.8)$$

Quá trình này đưa toàn bộ dữ liệu về một thang đo thống nhất, giảm thiểu các sai số do khác biệt sinh lý cá nhân.

- Chọn lọc đặc trưng đệ quy (Recursive Feature Elimination - RFE): Hệ thống áp dụng kỹ thuật RFE để loại bỏ các chỉ số có trọng số thấp, qua đó giữ lại bốn đặc trưng có giá trị phân loại tốt nhất giữa các nhóm đối tượng [144].
- Tinh chỉnh cấu hình thuật toán: Mô hình phân loại SVM (hạt nhân RBF) [145] được thiết lập để xác định biên quyết định phi tuyến. Cấu hình của thuật toán này được thử nghiệm và điều chỉnh lặp lại nhiều lần trên bộ dữ liệu mẫu nhằm tìm ra bộ thông số hoạt động hiệu quả nhất.
- Đánh giá hiệu năng tổng thể: Để kiểm tra năng lực thực tế, luận án áp dụng phương pháp kiểm chứng chéo LOOCV [146]. Tại mỗi vòng



Hình 3.10: Sơ đồ luồng đánh giá thực nghiệm và tối ưu hóa mô hình phân loại RLPTK

lập, dữ liệu của một trẻ được tách riêng để làm bài kiểm thử độc lập, trong khi mô hình sẽ học tập trên dữ liệu của toàn bộ $N - 1$ trẻ còn lại. Cách thực hiện này giúp ngăn chặn tình trạng mô hình ghi nhớ thói quen cá nhân của trẻ thay vì nhận diện đúng các dấu hiệu đặc trưng.

Sau khi hoàn tất quy trình xử lý và thực nghiệm, mô hình xác lập được bộ tham số phù hợp, hỗ trợ khả năng tổng quát hóa khi đưa ra các chỉ số phân loại tự động.

3.3.3.3 Kiểm chứng hiệu năng hỗ trợ phân loại và xác định dấu ấn sinh học

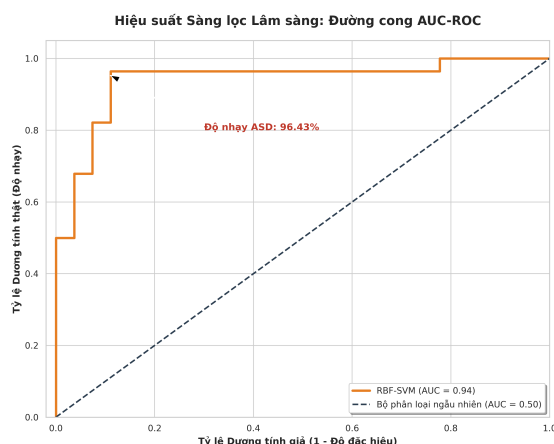
Kết quả thực nghiệm tại Bảng 3.1 cho thấy hệ thống đạt độ chính xác 92,73% (khoảng tin cậy 95% CI: 82,7% – 97,1%) và độ nhạy đạt 96,43%. Chỉ số này đáp ứng tiêu chuẩn hỗ trợ sàng lọc y tế trong việc phát hiện sớm nhóm trẻ RLPTK.

Bảng 3.1: Kết quả hiệu năng chẩn đoán của hệ thống đề xuất qua kiểm chứng chéo LOOCV

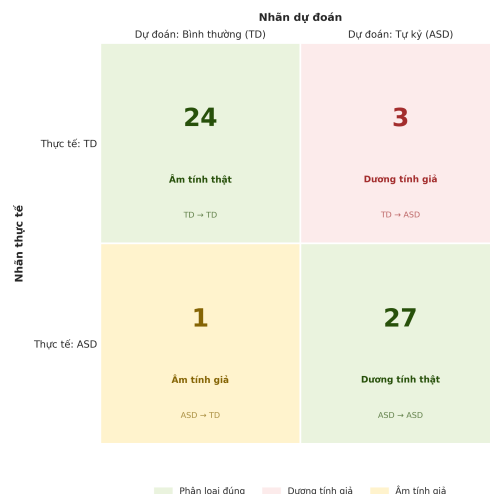
Chỉ số đánh giá	Giá trị	Khoảng tin cậy 95% (95% CI)
Độ chính xác tổng thể	92,73%	[82,7% - 97,1%]
Độ nhạy nhóm RLPTK	96,43%	–
Độ chuẩn xác nhóm RLPTK	90,00%	–
F1-Score	0,93	–
Chỉ số AUC	0,93	[0,88 - 0,98]

Hiệu năng phân tách trên các ngưỡng quyết định được minh họa qua đường cong ROC (Hình 3.11) và ma trận nhầm lẫn (Hình 3.12). Kết quả phân tích ma trận nhầm lẫn cho thấy ranh giới phân loại của mô hình SVM phân tách dữ liệu giữa nhóm trẻ RLPTK và nhóm đối chứng, đồng thời duy trì tỷ lệ dương tính giả ở mức thấp.

Để làm rõ cơ sở phân loại của mô hình, mức độ ảnh hưởng của các đặc trưng sinh trắc học được xác định thông qua biểu đồ ước lượng mật độ



Hình 3.11: Đường cong ROC của mô hình



Hình 3.12: Ma trận nhầm lẫn

hạt nhân (Kernel Density Estimation - KDE) kết hợp ma trận tương quan (Hình 3.13). Các đặc trưng chi phối chính bao gồm:

- Độ biến thiên chuyển động: Nhóm trẻ RLPTK có xu hướng dịch chuyển ảnh nhìn liên tục, tạo ra quỹ đạo mắt có tính biến thiên cao ($r = 1, 0$).
- Độ lệch chuẩn của tỷ lệ tập trung: Nhóm đối chứng hiển thị phân bố KDE biên độ hẹp và nhô cao, thể hiện sự hội tụ nhất quán trong hành vi duy trì tập trung ($r = -0, 50$).
- Mức độ tương tác mục tiêu: Phân bố dữ liệu giữa hai nhóm có sự chồng lấp lớn, thể hiện vai trò hỗ trợ thứ cấp trong phân loại ($r = -0, 34$).

Cuối cùng, chỉ số *Entropy quỹ đạo quét* được sử dụng để định lượng hành vi bao quát hình ảnh (Hình 3.14). Kết quả cho thấy nhóm đối chứng có quỹ đạo quan sát ổn định với mức độ phân tán thấp ($M = 0, 428, SD = 0, 122$). Ngược lại, nhóm trẻ RLPTK có mức entropy cao và phân tán không gian rộng ($M = 0, 537, SD = 0, 220$). Kết quả kiểm định Mann-Whitney U xác nhận sự khác biệt này có ý nghĩa thống kê ($p < 0, 05$), trong đó giá trị entropy của nhóm trẻ RLPTK cao hơn xấp xỉ 25%, phản ánh chiến lược tìm kiếm thị giác có tính phân tán cao.

3.3.3.4 Đánh giá đối chứng hiệu quả của phương pháp trích chọn đặc trưng

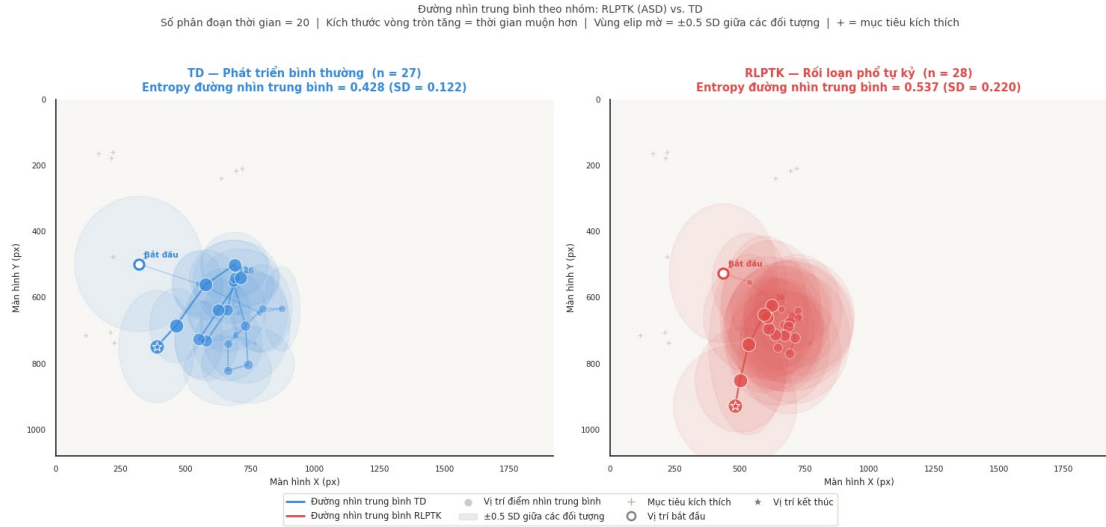
Nhằm định lượng mức độ cải thiện hiệu năng của phương pháp lai tích hợp thuật toán SAM và PeyeMMV so với kỹ thuật sử dụng dữ liệu nguyên



Hình 3.13: Phân tích độ tương quan và phân bố của các dấu ấn sinh học hành vi

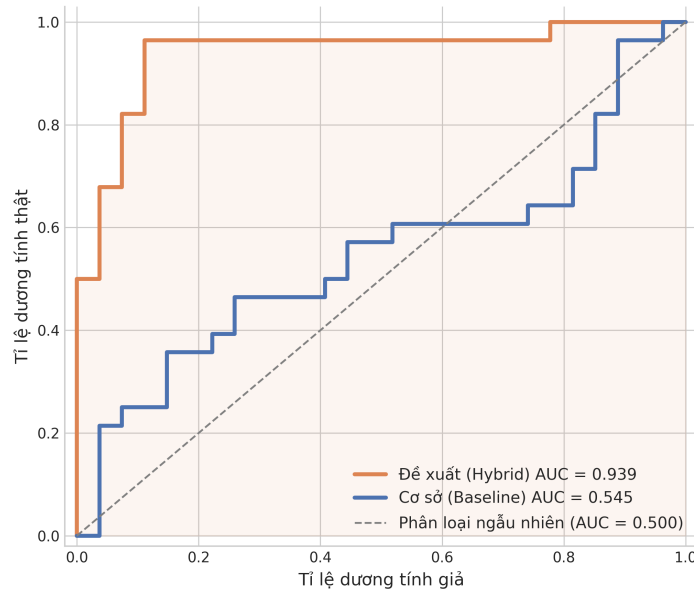
bản, một quy trình phân tích đối chứng đã được triển khai trên cùng một cấu hình mô hình SVM nhân RBF và chiến lược kiểm chứng chéo luân phiên từng cá thể LOOCV. Quá trình kiểm định này được thực thi trên hai hệ thống đầu vào bao gồm: (1) Cấu hình cơ sở sử dụng dữ liệu tọa độ nguyên bản đã được thiết lập tại Chương 2 và (2) Cấu hình lai đề xuất ứng dụng các đặc trưng cấu trúc ngữ nghĩa.

Sự khác biệt về năng lực phân loại tổng thể được minh họa định lượng thông qua phân tích đường cong ROC (Hình 3.15). Kết quả đánh giá chỉ ra



Hình 3.14: Đối sánh trực quan quỹ đạo đường quét trung bình nhóm.

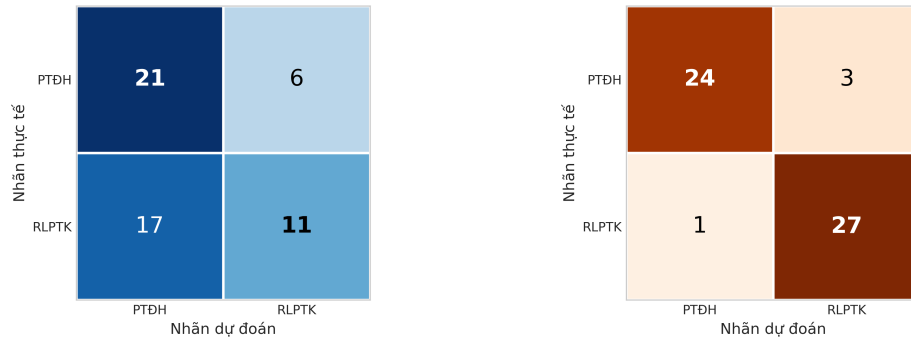
rằng đường cong ROC của mô hình lai dịch chuyển rõ rệt về phía góc trên bên trái, đi kèm với chỉ số AUC gia tăng đáng kể từ 0,545 lên 0,939. Số liệu thực nghiệm này phản ánh khả năng phân tách hiệu quả của hệ thống đề xuất trên nhiều ngưỡng quyết định độc lập.



Hình 3.15: Đường cong ROC đối chứng giữa cấu hình cơ sở và cấu hình lai đề xuất

Tiến hành phân tích chi tiết kết quả phân loại, Hình 3.16 biểu diễn ma trận nhầm lẫn của hai cấu hình phương pháp. Dữ liệu chỉ ra sự sụt giảm

đáng kể đối với tỷ lệ âm tính giả, đại diện cho các trường hợp hệ thống phân loại nhầm đối tượng mắc rối loạn phát triển thành nhóm phát triển điển hình. Cụ thể, trong khi cấu hình cơ sở ghi nhận 17 trường hợp nhận diện sai, việc ứng dụng phương pháp lai đã kiểm soát lượng sai số này xuống mức 1 trường hợp.

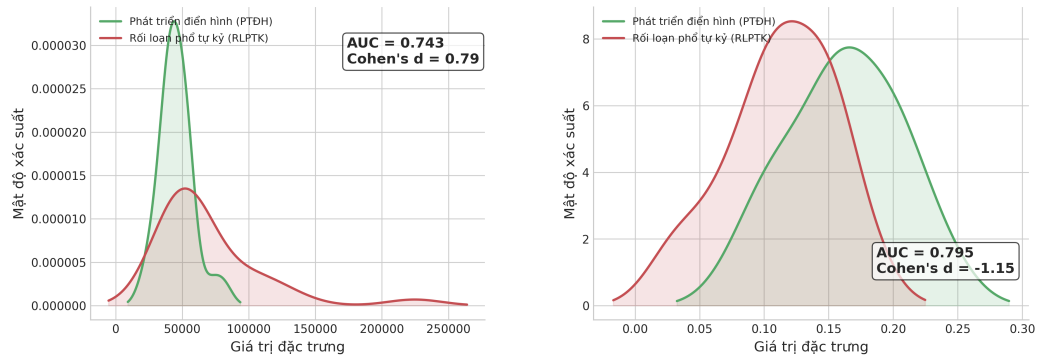


(a) Cấu hình cơ sở (Đặc trưng nguyên bản) (b) Cấu hình lai (Đặc trưng có ý nghĩa)

Hình 3.16: Đối sánh ma trận nhầm lẫn giữa hai cấu hình phương pháp

Sự chênh lệch về hiệu năng này bắt nguồn từ bản chất cấu trúc của tập hợp các đặc trưng biểu diễn. Hình 3.17 đối chiếu biểu đồ mật độ hạt nhân KDE đối với đặc trưng mang giá trị phân tách cao nhất tại từng cấu hình. Tại cấu hình cơ sở, đặc trưng không gian nguyên bản $Var(y)$ ghi nhận tỷ lệ chồng lấp phân bố cao giữa hai nhóm quan sát, làm suy giảm năng lực thiết lập ranh giới phân loại. Ngược lại, việc triển khai thuật toán SAM và PeyeMMV cho phép trích xuất các đặc trưng hành vi định hướng mục tiêu, diễn hình như *độ lệch chuẩn của tỷ lệ tập trung*. Hệ thống đặc trưng được ánh xạ này gia tăng khoảng cách phân bố giữa hai bộ dữ liệu, cung cấp cơ sở không gian hình học thuận lợi để mặt phẳng của mô hình thiết lập ranh giới phân tách.

Tổng hợp kết quả được trình bày trong Bảng 3.2, quá trình phân tích đối chứng chỉ ra rằng tiến trình chuyển đổi từ dữ liệu nguyên bản sang tập hợp các đặc trưng cấu trúc, thông qua phương pháp lai tích hợp cơ chế luồng quang học, hỗ trợ hiệu quả việc kiểm soát nhiễu động học. Cấu hình thuật toán được thiết lập đã cải thiện chỉ số độ nhạy từ 39,3% lên 96,4%, qua đó cung cấp nền tảng kỹ thuật tương thích với các tiêu chuẩn cơ bản của một hệ thống hỗ trợ đánh giá chẩn đoán.



(a) Phân bố đặc trưng không gian nguyên bản $\text{Var}(y)$ (b) Phân bố đặc trưng thị giác ngữ nghĩa

Hình 3.17: Phân bố mật độ hạt nhân KDE của các đặc trưng phân tách tốt nhất

Bảng 3.2: Đánh giá đối chứng hiệu năng phân loại giữa cấu hình cơ sở và cấu hình lai

Cấu hình phương pháp	Độ chính xác	Độ nhạy	Độ đặc hiệu	AUC	Âm tính giả
Cấu hình cơ sở (Mô hình nền và Dữ liệu thô)	58,2%	39,3%	77,8%	0,545	17/28
Cấu hình lai đề xuất (Tích hợp SAM và PeyeMMV)	92,7%	96,4%	88,9%	0,939	1/28

3.3.3.5 Đối sánh với các nghiên cứu liên quan

Để kiểm chứng năng lực hoạt động, hiệu năng của phương pháp trích xuất và phân loại đề xuất được đối chiếu với các công trình nghiên cứu đang được sử dụng phổ biến hiện nay. Bảng 3.3 tổng hợp kết quả đối sánh dựa trên các tiêu chí về độ chính xác và năng lực xử lý dữ liệu.

Kết quả cho thấy phương pháp đề xuất đạt độ chính xác 92,73% trong môi trường kích thích bằng video động. Đặt trong tương quan với các nghiên cứu sử dụng dữ liệu ảnh tĩnh (môi trường có mức độ nhiễu động học thấp hơn), tỷ lệ này phản ánh khả năng duy trì tính ổn định của hệ thống khi xử lý dữ liệu động nhằm hỗ trợ phát hiện các dấu ấn sinh học nhận thức. Việc duy trì được độ chính xác trong bối cảnh thực nghiệm, đặc biệt khi xuất hiện các cử động vi mô từ trẻ, cho thấy tính khả thi của giải pháp tự động hóa trong quá trình thu thập và bảo toàn độ tin cậy cho luồng dữ liệu điểm nhìn.

Bảng 3.3: Đối sánh hiệu năng với các nghiên cứu phân loại RLPTK tiên tiến

Nghiên cứu	Năm	Phương pháp	Loại kích thích	Accuracy	AUC
[13].	2024	CNN + Điểm nhìn	Ảnh tĩnh	92,00%	0,94
[147].	2021	ML + VR	Môi trường ảo	86,00%	-
[148].	2021	SVM + Điểm nhìn	Video xã hội	91,50%	0,92
[149].	2019	Random Forest	Video xã hội	88,00%	0,90
Luận án	2026	SVM + Spatial Entropy	Video động	92,73%	0,93

3.3.4 Đánh giá đặc trưng thị giác và cơ chế nhận thức ở trẻ rối loạn tính toán

Khác với các đặc trưng suy giảm tương tác xã hội ở nhóm trẻ rối loạn phổ tự kỉ, rào cản nhận thức chính của trẻ rối loạn tính toán nằm ở quá trình xử lý thông tin không gian, số học và giới hạn của bộ nhớ làm việc. Do đó, quy trình đánh giá kết quả thực nghiệm đối với nhóm trẻ này được thiết kế nhằm phân tích độc lập các thành phần nhận thức, bao gồm ba nội dung chính: (1) Phân tích thời lượng quan sát và ma trận chuyển đổi ánh nhìn nhằm định lượng hóa chiến lược xử lý thông tin và xác định các điểm gián đoạn trong quá trình đối chiếu dữ liệu; (2) Phân tích các đặc trưng nhận thức thuộc ba khía cạnh (thị giác, thính giác, không gian) nhằm đánh giá năng lực giải quyết các tác vụ tương ứng; và (3) Phân tích tương quan đa nguồn dữ liệu (chuyển động mắt, âm thanh, thao tác chuột) nhằm đánh giá mức độ đồng bộ giữa quá trình tiếp nhận giác quan và hành vi phản hồi vận động.

3.3.4.1 Phân tích thời lượng quan sát và ma trận chuyển đổi ánh nhìn

Giao diện bài tập được phân thành hai vùng quan sát chính: AOI-Đề_bài (chứa thông tin câu hỏi) và AOI-Đáp_án (chứa các phương án lựa chọn) (Hình 3.18). Kết quả phân tích dữ liệu theo dõi chuyển động mắt cho thấy sự khác biệt đáng kể trong chiến lược phân bổ chú ý giữa hai nhóm



Hình 3.18: Bố cục chính trang bài tập.

trẻ. Cụ thể, nhóm trẻ RLTT dành thời gian quan sát trung bình tại vùng AOI-Đề_bài là 10,82 giây, cao hơn so với nhóm đối chứng (6,02 giây). Ngược lại, tại vùng AOI-Đáp_án, nhóm đối chứng dành trung bình 12,99 giây, cao hơn đáng kể so với nhóm RLTT (7,11 giây).

Để làm rõ hơn cơ chế đằng sau sự khác biệt về thời gian quan sát nêu trên, luận án phân tích ma trận chuyển đổi ánh nhìn giữa hai vùng quan sát. Ma trận này mô tả tần suất ánh nhìn di chuyển giữa hai vùng AOI-Đề_bài và AOI-Đáp_án, gồm bốn phần tử với ý nghĩa cụ thể như sau:

- $m_{1,1}$: tần suất ánh nhìn di chuyển từ AOI-Đề_bài sang chính AOI-Đề_bài.
- $m_{1,2}$: tần suất ánh nhìn di chuyển từ AOI-Đề_bài sang AOI-Đáp_án, phản ánh hành vi chuyển từ đọc đề sang xem xét các phương án trả lời.
- $m_{2,1}$: tần suất ánh nhìn di chuyển ngược từ AOI-Đáp_án về AOI-Đề_bài, phản ánh hành vi quay lại đối chiếu và kiểm chứng thông tin đề bài sau khi đã xem đáp án.
- $m_{2,2}$: tần suất ánh nhìn di chuyển từ AOI-Đáp_án sang chính AOI-Đáp_án.

Trong đó, $m_{2,1}$ là chỉ số trọng tâm của phân tích này, vì đây là chỉ số trực tiếp phản ánh mức độ trẻ quay lại kiểm tra đề bài trước khi đưa ra lựa

chọn cuối cùng. Bảng 3.4 trình bày chi tiết các giá trị cận trên, cận dưới và trung bình nhìn của từng phần tử trong ma trận chuyển đổi ánh nhìn đối với mỗi nhóm.

Bảng 3.4: Giá trị cận trên, cận dưới và trung bình của ma trận chuyển đổi ánh nhìn theo nhóm

Nhóm	$m_{1,1}$	$m_{1,2}$	$m_{2,1}$	$m_{2,2}$
Đối chứng	0; 0; 0	63; 0; 8,44	64; 0; 8,56	0; 0; 0
Rối loạn tính toán	0; 0; 0	21; 0; 3,36	21; 0; 3,76	0; 0; 0

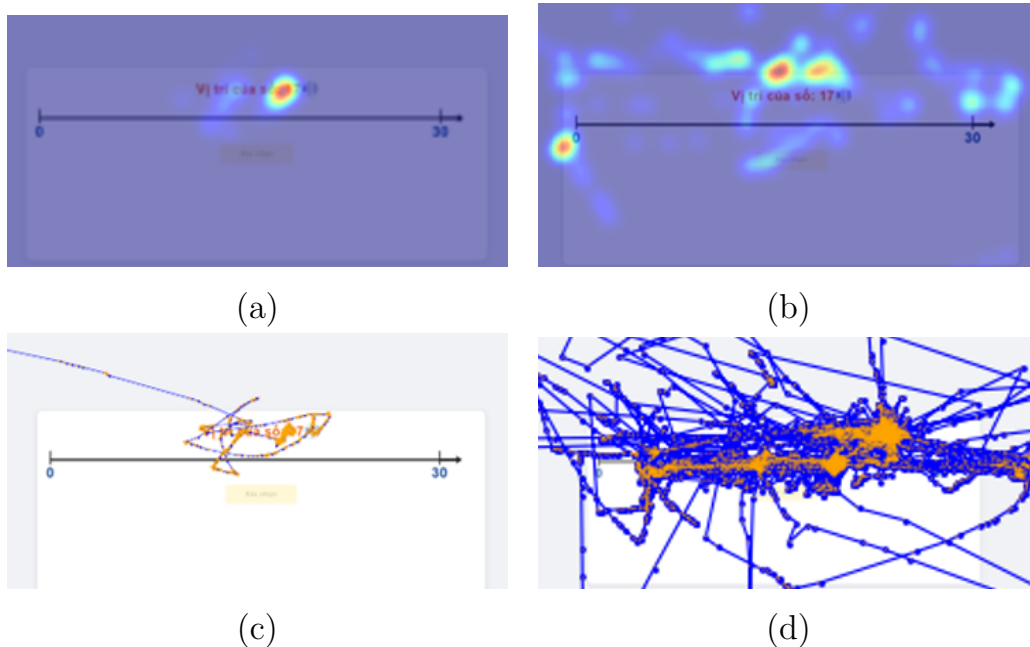
Kết quả cho thấy giá trị $m_{2,1}$ trung bình của nhóm đối chứng đạt 8,56 lượt, cao hơn khoảng 2,3 lần so với nhóm trẻ RLTT (3,76 lượt). Điều này cho thấy trẻ ở nhóm đối chứng có tần suất quay lại đối chiếu đề bài sau khi xem đáp án cao hơn rõ rệt so với nhóm RLTT.

Kết hợp với kết quả thời gian quan sát đã trình bày (10,82s so với 6,02s tại AOI-Đề bài; 7,11s so với 12,99s tại AOI-Đáp án), có thể nhận thấy sự khác biệt trong cách xử lý thông tin của hai nhóm: nhóm đối chứng tuy đọc đề bài trong thời gian ngắn hơn, nhưng bù lại liên tục quay lại đối chiếu đề bài trong suốt quá trình cân nhắc đáp án ($m_{2,1}$ cao), cho thấy trẻ không chỉ đọc đề một lần mà chủ động kiểm tra lại thông tin nhiều lần trước khi ra quyết định cuối cùng. Ngược lại, nhóm trẻ RLTT tuy dành nhiều thời gian hơn để đọc đề bài ngay từ đầu, nhưng tần suất quay lại kiểm tra đề bài sau khi đã chuyển sang vùng đáp án lại thấp hơn rõ rệt, cho thấy trẻ có xu hướng xử lý đề bài và đáp án như hai giai đoạn tách biệt, thiếu sự liên kết và đối chiếu qua lại giữa hai nguồn thông tin. Nói cách khác, việc đọc đề lâu hơn ở nhóm RLTT không đồng nghĩa với việc thông tin đó được duy trì và sử dụng hiệu quả trong quá trình ra quyết định, mà ngược lại, có thể phản ánh khó khăn trong việc mã hóa và ghi nhớ tạm thời nội dung đề bài, khiến trẻ phải đọc kỹ hơn nhưng lại không quay lại kiểm chứng được.

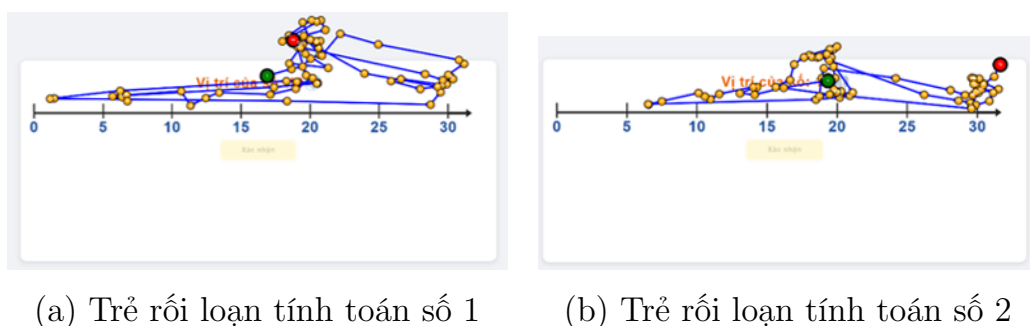
3.3.4.2 Phân tích đặc trưng nhận thức

Đặc trưng nhận thức của nhóm trẻ RLTT được đánh giá qua ba khía cạnh chính: nhận thức thị giác, nhận thức thính giác và nhận thức không gian.

Đặc trưng nhận thức thị giác: Thể hiện qua nhiệm vụ trên đường số (Hình 3.19), nhóm trẻ PTĐH có quỹ đạo nhìn tập trung và rõ ràng. Ngược lại, ánh nhìn của nhóm trẻ RLTT thường rải rác. Trong nhiệm vụ xác định vị trí số 13 (Hình 3.20), nhóm trẻ này thường bắt đầu ước lượng từ các điểm biên thay vì dựa trên các mốc tham chiếu gần, cho thấy hạn chế trong việc hình thành khái niệm không gian số học.

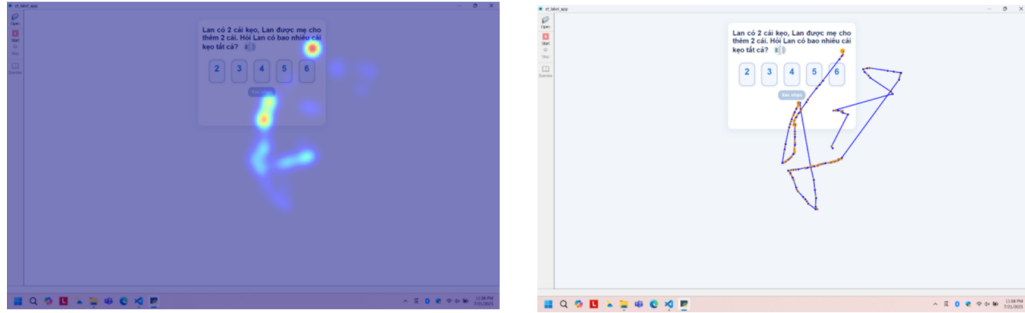


Hình 3.19: Bản đồ nhiệt và đường quét biểu diễn hành vi quan sát trong bài toán ước lượng trên trục số.



Hình 3.20: Đường quét cho các chiến lược ước tính khác nhau của trẻ rối loạn tính toán

Đặc trưng nhận thức thính giác: Thể hiện qua nhiệm vụ giải toán có lời văn khi nghe thông tin đề bài và chọn đáp án (Hình 3.21), nhóm trẻ PTĐH có xu hướng định hướng ánh nhìn vào các giá trị liên quan ngay trong quá trình tiếp nhận thông tin âm thanh của đề bài. Nhóm trẻ RLTT ít thể hiện hành vi định hướng này, dẫn đến các lựa chọn mang tính phân tán cao hơn.

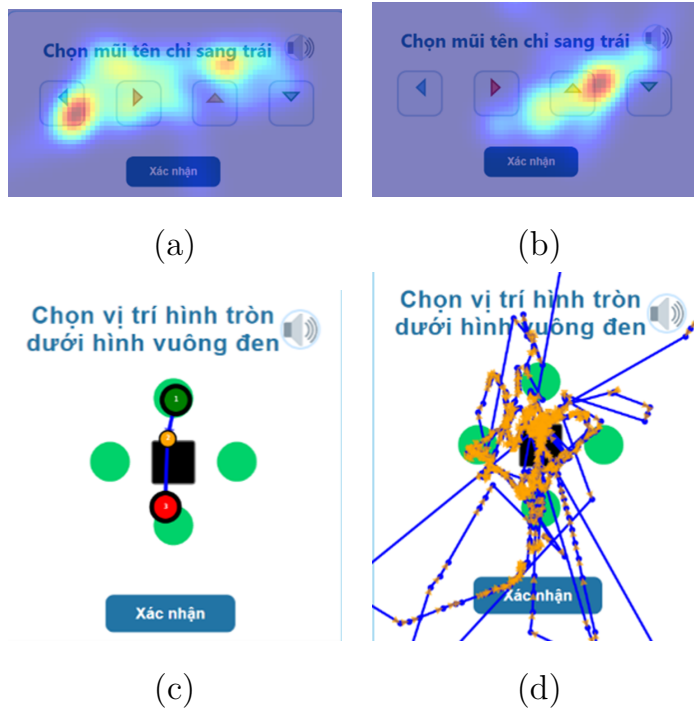


(a) Bản đồ nhiệt

(b) Đường quét

Hình 3.21: Bản đồ nhiệt và đường quét của trẻ rối loạn tính toán trong bài toán nghe và lựa chọn đáp án

Đặc trưng nhận thức không gian: Qua dữ liệu theo dõi hướng nhìn (Hình 3.22), nhóm đối chứng xác định mục tiêu với độ chính xác cao. Trong khi đó, trẻ RLTT thể hiện sự thiếu ổn định trong cách quan sát, thể hiện qua việc các đường quét mắt đan chéo nhau liên tục và không tuân theo một trật tự cụ thể.



(a)

(b)

(c)

(d)

Hình 3.22: Bản đồ nhiệt và đường quét trên các dữ liệu liên quan đến xử lý không gian thị giác

3.3.4.3 Phân tích kết hợp nhiều nguồn dữ liệu và mối liên hệ hành vi

Để phân tích mối liên hệ giữa các chuỗi hành vi của trẻ, luận án thực hiện đối chiếu dữ liệu chuyển động mắt, âm thanh và thao tác chuột. Kết quả phân tích tương quan Pearson xác định mối liên kết giữa thời lượng cố định mắt tại vùng AOI-ĐỀ_bài và độ trễ thao tác chuột có sự khác biệt giữa hai nhóm đối tượng:

- Nhóm trẻ PTDH: Hệ số tương quan đạt $r = 0,78$ ($p < 0,01$), thể hiện tính đồng bộ giữa quá trình tiếp nhận thông tin thị giác và hành vi vận động.
- Nhóm trẻ RLTT: Hệ số tương quan đạt $r = 0,31$ ($p < 0,05$), thể hiện mức độ liên kết thấp giữa hai tiến trình này.

Việc tích hợp dữ liệu âm thanh hỗ trợ phân tách khoảng thời gian nghe hướng dẫn và khoảng thời gian xử lý hình ảnh. Dữ liệu đối khớp nhãn thời gian cho thấy độ trễ phản hồi ở nhóm trẻ RLTT tập trung tại giai đoạn chuyển đổi thông tin, tính từ khi kết thúc tín hiệu âm thanh đến khi ánh nhìn hội tụ tại vùng AOI-Đáp_án.

Kết quả này chỉ ra rào cản của trẻ xuất phát từ tiến trình xử lý đồng bộ thông tin thính giác và thị giác, không do tốc độ vận động cơ học. Do đó, việc khai thác dữ liệu đa phương thức cung cấp thông tin về tiến trình nhận thức so với phương pháp sử dụng đơn lẻ dữ liệu chuyển động mắt.

3.4 Phương pháp tăng cường dữ liệu chuyển động mắt dựa trên tối ưu tiến hóa

3.4.1 Bài toán tăng cường dữ liệu chuyển động mắt

Việc tăng cường dữ liệu hành vi thị giác là hướng tiếp cận cần thiết để mở rộng quy mô tập huấn luyện. Tuy nhiên, quá trình này đặt ra một ràng buộc kỹ thuật khắt khe: hệ thống phải cân bằng giữa việc duy trì tính chân thực của sinh lý học nhãn cầu và việc thỏa mãn các điều kiện phân tích của bộ lọc PeyMMV. Các mô hình học sâu hiện tại thường vi phạm ngưỡng dung sai không gian vi mô, dẫn đến chuỗi tọa độ được tạo ra bị thuật toán nhận diện từ chối. Do đó, luận án xây dựng một khung tổng hợp dữ liệu

(Hình 3.23) vận hành theo nguyên lý tối ưu hóa tiến hóa (GA), tích hợp trực tiếp PeyeMMV thành một hàm đánh giá.

3.4.2 Mô hình hóa động lực học bằng chuyển động Brownian phân số (fBm)

3.4.2.1 Đặc tính sinh lý của các vi chuyển động mắt

Nhãn cầu duy trì dao động liên tục ngay cả khi tập trung vào mục tiêu tĩnh. Các dao động này, gọi là vi chuyển động mắt (Fixation Eye Movements - FEM), ngăn chặn hiện tượng thích nghi thần kinh. Để tăng cường dữ liệu chuyển động mắt, mô hình cần tích hợp:

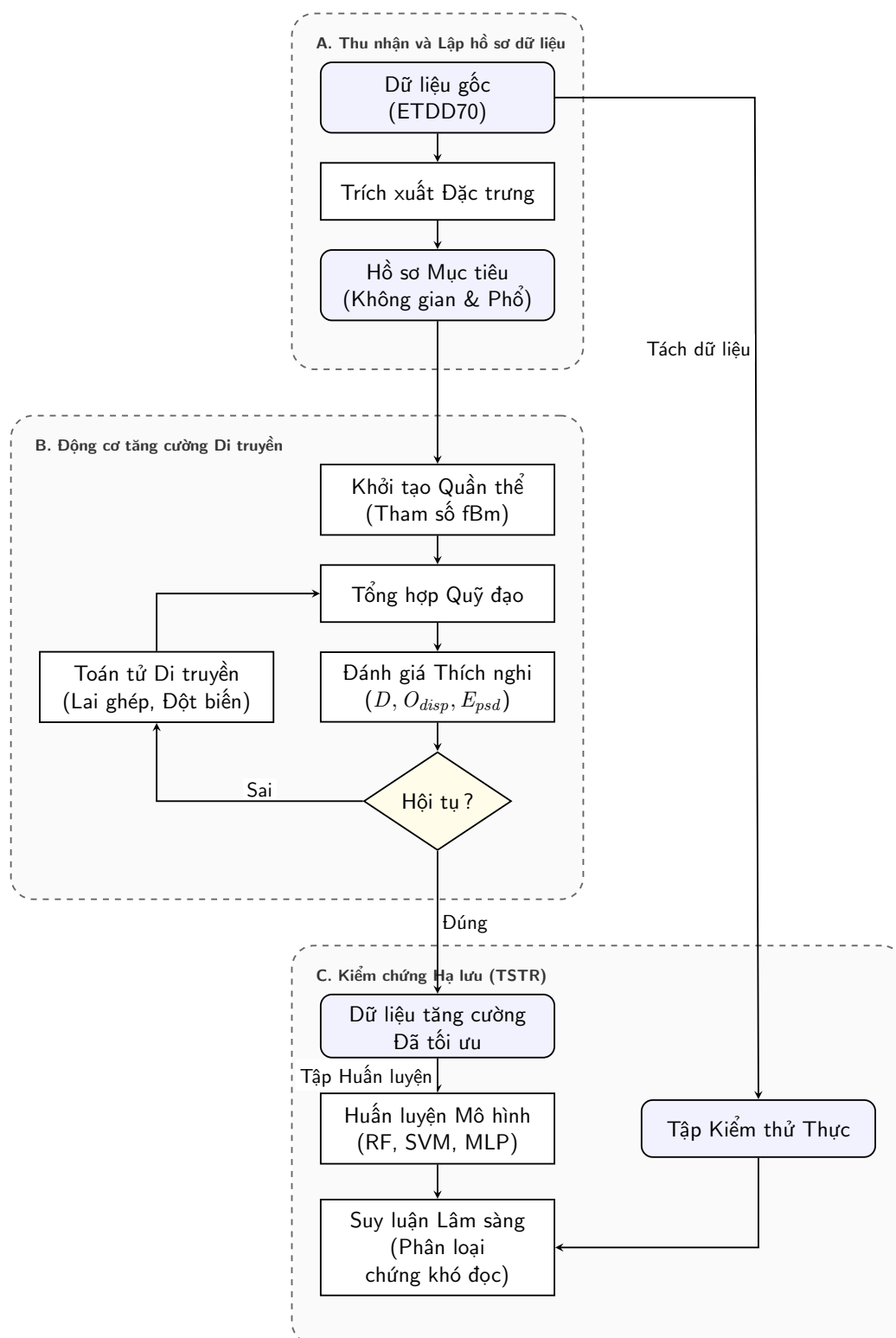
- Vi rung nhãn cầu: Các dao động tần số cao (30-100 Hz) với biên độ nhỏ.
- Độ trôi nhãn cầu: Chuyển động chậm giữa các vi chuyển động giật, thường gây ra sự phân tán không gian.
- Vi chuyển động giật: Các chuyển động nhanh có chức năng tái định vị tiêu điểm.

Chuỗi tín hiệu chuyển động của nhãn cầu mang đặc trưng của phổ nhiễu hồng $1/f$. Khác với tính ngẫu nhiên độc lập của nhiễu trắng, cấu trúc nhiễu hồng sở hữu tính tự tương quan giữa các mẫu thời gian kế tiếp. Việc mô phỏng chính xác đặc tính phổ này là một yêu cầu kỹ thuật nhằm duy trì sự tương đồng về đặc tính sinh lý trong quá trình tăng cường dữ liệu quỹ đạo.

3.4.2.2 Mô hình hóa đặc tính động học nhãn cầu dựa trên fBm

Để mô phỏng đặc tính nhiễu hồng $1/f$, quá trình trôi nhãn cầu được mô hình hóa thông qua chuyển động Brownian phân số (fBm). Tham số Hurst (H) quyết định mức độ tự tương quan của chuỗi thời gian [150]:

- $H = 0,5$: Quá trình tương đương với chuyển động Brownian tiêu chuẩn.
- $H > 0,5$: Chuỗi có tính bền vững, duy trì xu hướng theo thời gian.
- $H < 0,5$: Chuỗi có tính đối nghịch, đặc trưng cho các vi dao động tần số cao.



Hình 3.23: Kiến trúc hệ thống tăng cường dữ liệu di truyền và kiểm chứng TSTR.

3.4.2.3 Kiến trúc tổng thể của hệ thống

Quy trình vận hành được cấu trúc thành chu trình khép kín (Hình 3.23), bao gồm ba phân hệ chính:

1. Phân hệ thu nhận và trích xuất đặc trưng: Xây dựng tham số đường cơ sở từ bộ dữ liệu thực (ETDD70).
2. Phân hệ tăng cường di truyền: Khởi tạo quần thể tham số và ánh xạ thành quỹ đạo nhìn qua mô hình fBm; bộ lọc PeyeMMV được sử dụng làm hàm mục tiêu trong vòng lặp tối ưu hóa.
3. Phân hệ đánh giá thực nghiệm qua giao thức huấn luyện tăng cường, kiểm tra thực tế (Train Synthetic Test Real - TSTR): Sử dụng dữ liệu tăng cường để huấn luyện các mô hình học máy và đánh giá khả năng khái quát hóa trên bộ dữ liệu thực.

3.4.2.4 Cấu trúc hàm đánh giá dựa trên thuật toán PeyeMMV

Việc ánh xạ ngược cho hệ thống được mô hình hóa thành bài toán tối ưu hóa có ràng buộc:

$$\text{Minimize } \mathcal{J}(G) = \mathbb{I}(F(G) \neq f_{\text{target}}) + \beta \cdot \text{KL}(P_G \mid P_{1/f}) \quad (3.9)$$

Trong đó:

- $\mathbb{I}$ là hàm chỉ báo (phạt khi mô hình nhận diện sai).
- $\text{KL}(P_G \parallel P_{1/f})$ là phân kỳ Kullback-Leibler giữa mật độ phổ công suất (Power Spectral Density - PSD) của quỹ đạo G và nhiễu hồng sinh lý $1/f$.

3.4.2.5 Thiết kế hàm thích nghi

Bài toán tối ưu hóa đa mục tiêu được chuyển đổi thành một hàm mục tiêu vô hướng thông qua phương pháp tổng có trọng số:

$$\mathcal{F}(G) = w_1 \cdot D - w_2 \cdot O - w_3 \cdot E \quad (3.10)$$

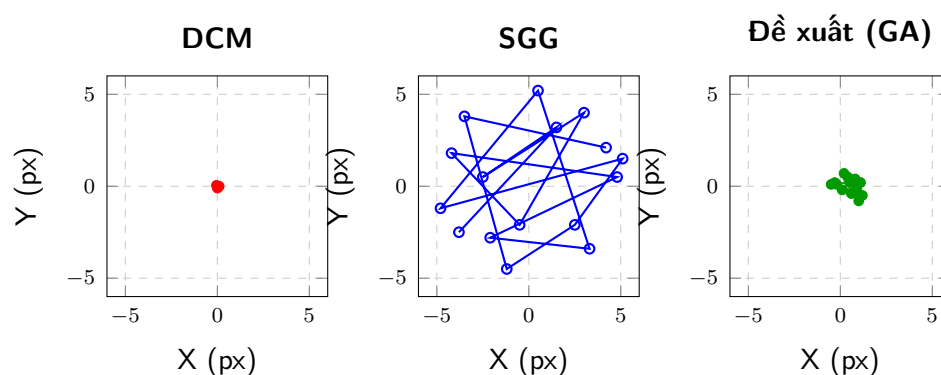
Trong đó, D đại diện cho tỷ lệ phát hiện, O đại diện cho tỷ lệ điểm ngoại lai, và E đại diện cho mức độ sai số về phổ và không gian. Tham số trọng số w_3 được tinh chỉnh thông qua kỹ thuật tìm kiếm dạng lưới nhằm ưu tiên duy trì sự tương đồng đối với phổ nhiễu hồng $1/f$. Cơ chế này giúp kiểm

soát xu hướng hội tụ của dữ liệu quỹ đạo được tăng cường, qua đó bảo đảm chuỗi dữ liệu đầu ra đáp ứng các đặc tính sinh lý tự nhiên của nhãn cầu.

3.5 Thực nghiệm và đánh giá hiệu quả của phương pháp tăng cường dữ liệu

3.5.1 Đánh giá tỉ lệ phát hiện và độ phân tán không gian

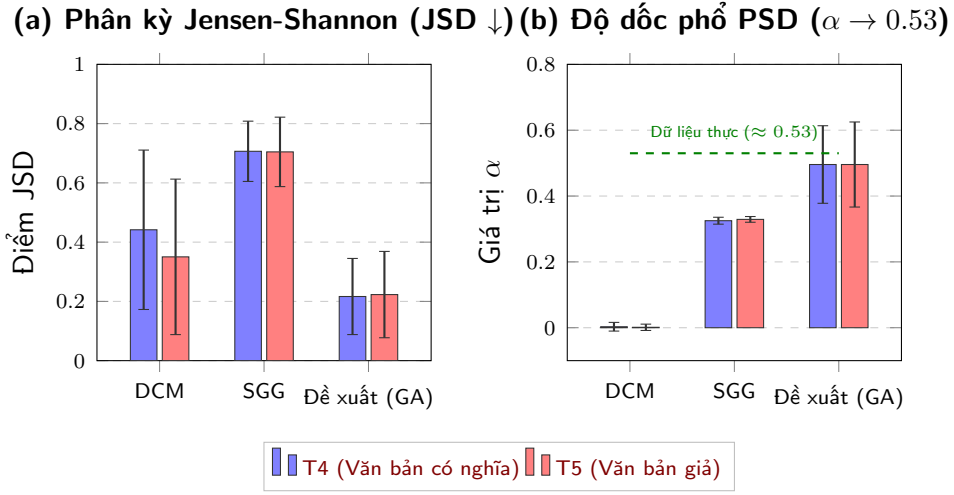
Để đánh giá hiệu năng, phương pháp đề xuất (GA) được đối sánh với hai phương pháp cơ sở mang đặc tính đối lập: Sinh ngẫu nhiên Gaussian (Stochastic Gaussian Generation - SGG) và cực tiểu hóa tâm xác định (Deterministic Centroid Minimization - DCM). Dựa trên phân tích phân bố không gian tại Hình 3.24, hai phương pháp cơ sở đều bộc lộ giới hạn rõ rệt. Cụ thể, phương pháp DCM loại bỏ hoàn toàn yếu tố ngẫu nhiên, dẫn đến việc tạo ra một cụm điểm tĩnh với phương sai xấp xỉ 0. Đặc điểm này khiến hệ thống kém hiệu quả trong việc mô phỏng cấu trúc FEM. Ở thái cực ngược lại, phương pháp SGG tạo ra độ phân tán nhiều trên diện rộng, vi phạm ngưỡng dung sai và không thỏa mãn các điều kiện xác định điểm cố định của thuật toán PeyreMMV. Khắc phục hai nhược điểm này, phương pháp GA thiết lập một điểm cân bằng: hệ thống vừa tạo ra mức phương sai đủ để tái hiện hiện tượng vi rung nhãn cầu, vừa kiểm soát chặt chẽ chuỗi tọa độ bên trong biên độ không gian mục tiêu.



Hình 3.24: Phân tích đối sánh độ phân tán không gian của các cụm điểm cố định tăng cường. Đỏ (DCM - Tĩnh hoàn toàn, mất phương sai sinh lý), Xanh dương (SGG - Nhiều trắng không cấu trúc, dễ bị bộ lọc từ chối), và Xanh lá (GA - Phân bố sinh lý tự nhiên, duy trì phổ $1/f$).

3.5.2 Đánh giá độ chân thực phổ và đặc tính nhiễu hồng

Phân tích định lượng về độ tương đồng phổ (Hình 3.25) cho thấy phương pháp GA đạt độ phân kỳ Jensen Shannon (Jensen Shannon Divergence - JSD) ở mức 0,22. Độ dốc PSD của thuật toán đạt $\alpha \approx 0,50$, tiệm cận với phân bố của dữ liệu thực nghiệm ($\alpha \approx 0,53$). Kết quả này khẳng định khả năng của hệ thống trong việc tái tạo cấu trúc nhiễu hồng, qua đó duy trì sự tương đồng về đặc tính sinh lý trong chuyển động của nhãn cầu.



Hình 3.25: Kiểm chứng thống kê về các chỉ số độ chân thực phổ. (a) GA đạt phân kỳ cực thấp. (b) GA tái tạo thành công động lực nhiễu hồng $1/f$.

3.5.3 Khả năng bảo tồn các đặc trưng nhận thức

Mục tiêu của việc tăng cường dữ liệu là bảo tồn các dấu ấn hành vi đặc trưng. Như đã thiết lập tại mục 3.2.3 (Chương 3), tỉ lệ bước nhảy mắt hồi quy (R_{reg}) đóng vai trò là một tham số quan trọng trong phân tích hành vi đọc. Để kiểm định khả năng này, chỉ số R_{reg} được trích xuất từ bộ dữ liệu tăng cường và đối chiếu trực tiếp với bộ dữ liệu thực nghiệm (ETDD70):

- Bộ dữ liệu thực: Chỉ số R_{reg} trung bình của nhóm trẻ mắc chứng khó đọc đạt 32,4%, trong khi nhóm phát triển điển hình đạt 18,7%.
- Bộ dữ liệu tăng cường: Chỉ số R_{reg} trung bình của nhóm trẻ mắc chứng khó đọc đạt 31,1%, trong khi nhóm phát triển điển hình đạt 19,2%.

Mức độ sai lệch về phân bố đặc trưng giữa hai nguồn dữ liệu được kiểm soát dưới 5%. Kết quả này cho thấy khung phương pháp dựa trên chuyển

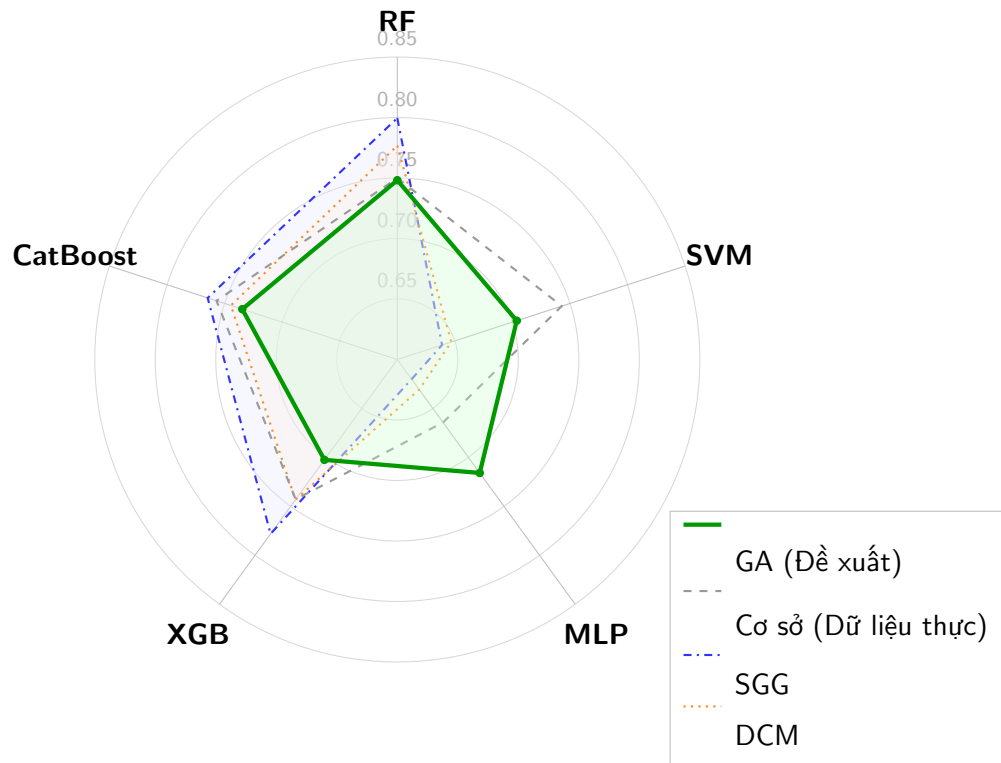
động Brownian phân số (fBm) có khả năng bảo toàn cấu trúc hành vi đọc. Sự tương đồng về mặt thống kê chính là cơ sở xác nhận mức độ tin cậy của bộ dữ liệu tăng cường khi được sử dụng làm đầu vào huấn luyện cho các mô hình học máy.

3.5.4 Đánh giá hiệu năng phân loại chẩn đoán

	Cơ sở (Real)	SGG	DCM	Đề xuất (GA)
RF	0.7500	0.7996	0.7768	0.7482
SVM	0.7429	0.6389	0.6469	0.7038
MLP	0.6643	0.6201	0.6308	0.7159
XGB	0.7429	0.7782	0.7433	0.7025

Hình 3.26: Bản đồ nhiệt biểu diễn độ chính xác phân loại TSTR. Mapping cường độ màu sắc cô lập sự suy giảm nghiêm trọng của các phương pháp sinh dữ liệu không cấu trúc (SGG, DCM) trên các mô hình nhạy cảm ranh giới (SVM, MLP).

Thông qua giao thức TSTR, dữ liệu tăng cường được sử dụng để huấn luyện mô hình chẩn đoán chứng khó đọc, sau đó kiểm thử trên bộ dữ liệu thực (ETDD70). Phân tích Hình 3.26 cho thấy phản ứng của các mô hình SVM và mạng nơ-ron đa lớp (Multilayer Perceptron - MLP) đối với dữ liệu đầu vào. Các phương pháp SGG và DCM dẫn đến sự giảm độ chính xác của mô hình xuống mức 0,62 – 0,64 do làm thay đổi cấu trúc không gian dữ liệu. Ngược lại, dữ liệu từ phương pháp GA hỗ trợ tăng hiệu năng mô hình MLP từ 0,6643 (trên dữ liệu thực) lên 0,7159 và duy trì độ ổn định cho SVM ở mức 0,7038. Kiểm định ANOVA xác nhận sự khác biệt về hiệu năng giữa phương pháp GA và các phương pháp khác có ý nghĩa thống kê ($p < 0,05$). Biểu đồ Radar (Hình 3.27) thể hiện phạm vi bao phủ đồng đều của dữ liệu từ GA, đáp ứng các yêu cầu về bảo toàn ranh giới phân loại phi tuyến cho các tác vụ phân tích hành vi phức tạp.



Hình 3.27: Phân bố hiệu năng đa chiều. Đa giác của GA thể hiện hình thái bao phủ tròn đều, ôm sát quỹ đạo của dữ liệu cơ sở trên toàn bộ 5 kiến trúc thuật toán.

Kết chương

Chương này đã trình bày hai nhóm phương pháp xử lý dữ liệu chuyển động mắt nhằm thiết lập nguồn dữ liệu đầu vào cho các mô hình phân tích chẩn đoán.

Thứ nhất, luận án đề xuất phương pháp trích xuất đặc trưng thị giác tự động thông qua việc tích hợp mô hình phân đoạn SAM và thuật toán PeyeMMV. Cấu trúc này giúp hệ thống nhận diện và bám sát liên tục các vùng quan tâm trên chuỗi khung hình động, qua đó khắc phục những hạn chế của phương thức gán nhãn thủ công. Kết quả thực nghiệm kiểm chứng chéo cho thấy giải pháp này hỗ trợ giảm thiểu nhiễu động học, cải thiện độ chính xác trong bài toán phân loại nhóm trẻ rối loạn phổ tự kỉ đạt mức 92,73% và hạn chế đáng kể tỉ lệ âm tính giả.

Thứ hai, luận án đề xuất khung tăng cường dữ liệu quỹ đạo ánh nhìn dựa trên nguyên lý tối ưu hóa tiến hóa (GA) kết hợp mô hình chuyển động Brownian phân số (fBm) nhằm khắc phục giới hạn về quy mô bộ dữ liệu.

Cơ chế này chuyển đổi thuật toán xác định điểm cố định thành một hàm đánh giá để trực tiếp điều hướng tiến trình tăng cường dữ liệu. Kết quả là hệ thống đã sinh ra các chuỗi quỹ đạo duy trì được sự tương đồng về cấu trúc phổ nhiễu hồng $1/f$ và các đặc trưng hành vi đọc, cung cấp nguồn dữ liệu tăng cường cho quá trình huấn luyện mô hình học máy.

Hệ thống các phương pháp trích xuất đặc trưng và kỹ thuật tăng cường dữ liệu này đã thiết lập một cơ sở kỹ thuật thống nhất. Dữ liệu này đóng vai trò tiền đề trực tiếp phục vụ quá trình triển khai các mô hình phân tích và thiết kế giải pháp can thiệp ở chương tiếp theo.

Các kết quả nghiên cứu liên quan đến nội dung chương này đã được công bố trong các công trình sau:

a. Các công trình đã được công bố/đăng tải

- **Thi-Quynh-Hoa Nguyen**, Duy-Hai Nguyen, Thanh-Ha Le, and Thi-Duyen Ngo. An Eye-Tracking System for Extracting and Visualizing Visual Features of Dyscalculia in Children. In: *Proceedings of the 14th International Symposium on Information and Communication Technology (SoICT 2025)*, Nha Trang, Vietnam, 2025. DOI: https://doi.org/10.1007/978-981-92-2590-3_31
- **Thi-Quynh-Hoa Nguyen**, Duy-Hai Nguyen, Gia-Loc Tran, Thanh-Ha Le, and Thi-Duyen Ngo. Bridging the Data Gap: A Genetic Algorithm Approach Using Inverse Python Implementation of Eye Movement Metrics and Visualization Logic. *Applied Soft Computing*, 2026 (WoS/SCOPUS Q1, IF: 7.8). DOI: <https://doi.org/10.1016/j.asoc.2026.116328>
- Đơn đăng ký sáng chế đã được chấp nhận. Số đăng ký: 1-2024-07110. Tên sáng chế: “Phương pháp và hệ thống cung cấp thông tin về sự chú ý thị giác”
- Báo cáo chuyên đề tại Diễn đàn Châu Á lần thứ nhất về hỗ trợ trẻ rối loạn phổ tự kỉ và trẻ rối loạn phát triển khác. Địa điểm: Hà Nội, Việt Nam, 06/03/2026.

b. Các công trình đã nộp (Đang trong quá trình xét duyệt)

- **Thi-Quynh-Hoa Nguyen**, Duc-Duy Le, Dang-Khoa Ta, Nu-Tam-An Nguyen, Thi-Cam-Huong Nguyen, Thanh-Ha Le, and Thi-Duyen Ngo. A Zero-Shot Segmentation Framework Using Eye-Tracking and the Segment Anything Model: Automated Extraction of Visual Biomarkers in Autism Spectrum Disorder. Submitted to *Engineering Applications of Artificial Intelligence*, 2026 (**WoS/SCOPUS Q1**).

Chương 4

CÁC PHƯƠNG PHÁP CAN THIỆP CÁ NHÂN HÓA HỖ TRỢ TRẺ RỐI LOẠN PHÁT TRIỂN

Chương này trình bày các phương pháp hỗ trợ can thiệp cá nhân hóa luận án đề xuất cho trẻ rối loạn phát triển, tập trung vào hai hướng tiếp cận: thiết kế môi trường tương tác số nhằm giảm tải nhận thức và xây dựng các mô hình tính toán để cá nhân hóa lộ trình can thiệp dựa trên dữ liệu sinh trắc học. Kiến trúc hệ thống được tổ chức thành các phân hệ chức năng, trong đó mỗi phân hệ đảm nhiệm một khâu đánh giá riêng biệt trong quy trình thực nghiệm. Cụ thể, nhóm phân hệ tương tác được triển khai nhằm thu thập luồng dữ liệu hành vi thị giác, trong khi nhóm phân hệ thuật toán được kiểm chứng để thiết lập cơ sở cho việc xây dựng mô hình tích hợp ở các giai đoạn sau.

4.1 Giới thiệu

4.1.1 Bối cảnh và định hướng phương pháp luận

Sự đa dạng trong biểu hiện hành vi của trẻ rối loạn phát triển đặt ra yêu cầu cao về khả năng tùy biến đối với các hệ thống hỗ trợ. Việc tích hợp toàn bộ quy trình phân tích và can thiệp vào một kiến trúc phần mềm thường bộc lộ nhiều hạn chế về hiệu suất phân bổ tài nguyên, đồng thời làm gia tăng trạng thái tính toán [151, 152].

Để khắc phục các rào cản này và đáp ứng sự phân hóa về thiết bị cũng như mục tiêu đo lường, luận án đề xuất *Kiến trúc ra quyết định lai luồng kép (Dual-stream Hybrid Decision Architecture)*. Trong kiến trúc này, tiến trình xử lý dữ liệu đa nguồn được phân tách thành hai luồng suy diễn độc lập. Cấu trúc phân tán này không chỉ giúp giảm bớt gánh nặng tính toán cho hệ thống, mà còn đảm bảo khả năng tương tác linh hoạt với người sử dụng thông qua các cơ chế sau:

- Luồng 1: Phân tích hành vi thị giác tức thời: Khai thác đặc trưng sinh trắc học đưa vào một *hệ cơ sở luật* để đưa ra các phản hồi điều chỉnh kịp thời về mặt không gian vật lý và giao diện [24]. Luồng này đảm bảo tính minh bạch và khả năng can thiệp nhanh.
- Luồng 2: Mô hình hóa trạng thái tri thức: Khai thác dữ liệu hành vi tương tác trên phân hệ nền tảng số (Web và di động) làm đầu vào cho *mạng Bayes động (Dynamic Bayesian Network - DBN)*. Luồng này chịu trách nhiệm suy diễn mức độ thành thạo kỹ năng của trẻ để điều hướng bài tập [153].

Các tiến trình tính toán này được tích hợp và triển khai trên các nền tảng số chuyên biệt (như ứng dụng EzMath, EzRead), cung cấp giao diện thích ứng giúp giảm tải nhận thức và hạn chế yếu tố gây nhiễu cho trẻ [45, 154].

4.1.2 Động lực và đóng góp

Dựa trên nguyên lý triển khai độc lập và hai hướng nghiên cứu trọng tâm, các đóng góp của luận án bao gồm:

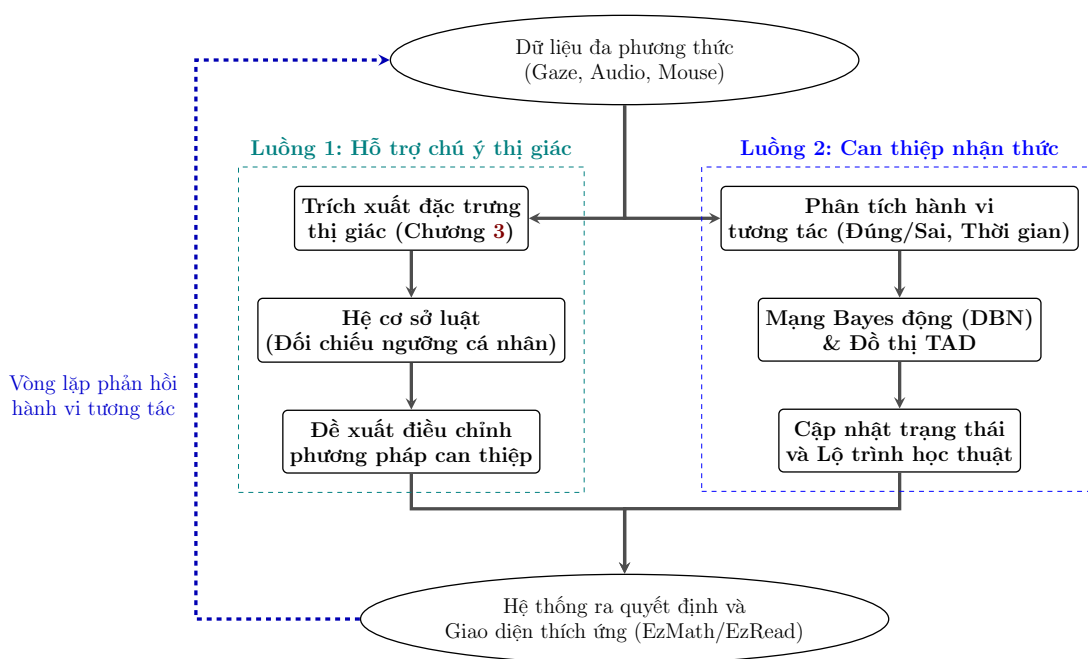
- Phát triển nền tảng web EzMath và ứng dụng di động EzRead dành cho trẻ mắc chứng khó đọc và rối loạn tính toán, tích hợp cấu trúc giao diện thích ứng nhằm giảm tải nhận thức dựa trên đặc điểm hành vi của trẻ.
- Xây dựng hệ cơ sở luật phân tích dữ liệu hành vi thị giác từ công nghệ theo dõi chuyển động mắt dành cho nhóm trẻ rối loạn phổ tự kỉ, trong đó mô hình vận hành dựa trên các ngưỡng tham số được cá nhân hóa nhằm cung cấp khuyến nghị hỗ trợ điều chỉnh can thiệp phù hợp.

- Xây dựng khung tính toán sử dụng DBN và đồ thị phân tích phụ thuộc nhiệm vụ (Task Analysis Diagram - TAD) để định lượng và suy diễn các trạng thái nhận thức của trẻ trong quá trình tương tác.

4.2 Kiến trúc hệ thống can thiệp vòng kín

4.2.1 Mô hình khái niệm và vai trò của dữ liệu chuyển động mắt

Luận án đề xuất mô hình khái niệm về hệ thống can thiệp vòng kín làm khung tham chiếu [155] nhằm kết hợp các nền tảng tương tác số và các mô hình tính toán thành một cấu trúc thống nhất. Sơ đồ kiến trúc tổng thể được minh họa tại Hình 4.1. Cấu trúc này thiết lập cơ chế trao đổi thông tin giữa phân hệ giao diện người dùng và phân hệ xử lý trung tâm.



Hình 4.1: Kiến trúc hệ thống can thiệp luồng kép phân tách luồng xử lý thị giác (cơ sở luật) và luồng nhận thức (mạng Bayes)

Trong mô hình này, dữ liệu theo dõi chuyển động mắt được sử dụng làm cơ sở phân tích trạng thái nhận thức của trẻ. Cụ thể, luồng dữ liệu thô sau khi trích xuất được đối chiếu với các ngưỡng tham số cá nhân hóa thông qua hệ cơ sở luật, từ đó chuyển đổi thành tín hiệu đầu vào phục vụ hai tiến trình: cung cấp các thông tin khuyến nghị nhằm hỗ trợ chuyên gia điều

chỉnh phương pháp can thiệp, đồng thời cập nhật tham số cho các mô hình tính toán nhận thức.

Theo nguyên lý triển khai độc lập, việc tích hợp chuỗi xử lý dữ liệu này vào một kiến trúc phần mềm được thay thế bằng phương pháp phân rã do sự khác biệt về hạ tầng thiết bị thực tế [156]. Do đó, luận án triển khai thực nghiệm hệ thống dưới dạng các cơ chế độc lập, bao gồm:

- Cơ chế đề xuất điều chỉnh can thiệp: Được cấu trúc theo hai phương án để tương thích với từng môi trường thực nghiệm. Đối với nhóm trẻ RLPTK, hệ cơ sở luật [157] thực hiện phân tích chuỗi dữ liệu theo dõi chuyển động mắt để cung cấp các thông tin đề xuất, hỗ trợ chuyên gia điều chỉnh phương pháp tương tác ở môi trường vật lý hoặc tùy biến các yếu tố phần mềm. Đối với nhóm trẻ rối loạn học tập, lộ trình can thiệp số hóa được cấu trúc trên các nền tảng web và di động, trong đó việc tinh chỉnh giao diện được xác định là một cơ chế thích ứng nhằm giảm tải nhận thức cho trẻ.
- Cơ chế cập nhật tham số mô hình qua dữ liệu hành vi: Được hình thức hóa toán học bằng cấu trúc mạng Bayes động [158]. Trong giai đoạn hiện tại, hệ thống thu thập các tham số tương tác học tập của trẻ để làm dữ liệu đầu vào cho mô hình mạng Bayes nhằm ước lượng trạng thái nhận thức, đồng thời là nền tảng lý thuyết cho việc tích hợp dữ liệu sinh trắc học ở các giai đoạn phát triển tiếp theo.

Mối tương quan giữa hành vi thị giác và phản hồi của hệ thống được tổng hợp tại Bảng 4.1. Các chỉ số ánh xạ này thể hiện vai trò của dữ liệu chuyển động mắt trong việc xác định trạng thái chuyển đổi ở các quy trình thực nghiệm và các mô hình toán học.

4.2.2 Cơ sở phương pháp thực nghiệm: Phương pháp đánh giá đa cấp độ và thiết kế thực nghiệm trường hợp đơn

Cơ chế đề xuất điều chỉnh can thiệp: Được cấu trúc để tối ưu hóa tiến trình học tập dựa trên đặc điểm nhận thức của trẻ thông qua các ứng dụng số hóa. Trong nghiên cứu này, việc đánh giá hiệu quả được triển khai thông qua hai phương thức tiếp cận tương hỗ: (1) Áp dụng thiết kế thực nghiệm trường hợp đơn (Single Case Experimental Design - SCED) thông qua việc can thiệp trên một trẻ cụ thể nhằm theo dõi sự thay đổi hành vi thị giác,

Bảng 4.1: Bảng tổng hợp: Cơ sở ánh xạ từ tín hiệu thị giác sang hành động thích ứng của hệ thống

Tín hiệu thị giác	Ước lượng trạng thái nhận thức	Hành động thích ứng của hệ thống
Tỷ lệ cố định thị giác tại một vùng trên 80%	Sự thu hẹp phạm vi chú ý và suy giảm tiến trình nhận thức ở trẻ RLPTK.	Điều chỉnh phân hệ tương tác: Làm mờ vùng nhiều và phát tín hiệu âm thanh điều hướng ánh nhìn.
Quý đạo ánh nhìn phân tán kết hợp số lượng điểm cố định cao	Tình trạng quá tải nhận thức ngoại lai do nhiều loạn thị giác.	Điều chỉnh phân hệ tính toán: Đề xuất giảm mức độ phức tạp của nhiệm vụ trên đồ thị.
Thời lượng cố định thị giác kéo dài	Tình trạng quá tải nhận thức nội tại do thiếu hụt kiến thức quan trọng.	Điều chỉnh phân hệ tính toán: Cập nhật xác suất trượt trong mạng Bayes động và kích hoạt truy vết ngược.

thao tác tương tác và mức độ giảm tải nhận thức theo thời gian; (2) Tiến hành khảo sát diện rộng đối với hơn 70 chuyên gia và giảng viên đang trực tiếp giảng dạy trong lĩnh vực giáo dục đặc biệt để thu thập đánh giá khách quan về tính sư phạm, tính thực tiễn và khả năng đáp ứng của hệ thống.

Cơ chế cập nhật tham số mô hình qua dữ liệu hành vi: Được hình thức hóa toán học bằng cấu trúc mạng Bayes động [158]. Hệ thống thu thập chuỗi tham số tương tác trong quá trình học tập của trẻ để cung cấp dữ liệu đầu vào cho mô hình mạng Bayes, từ đó đánh giá và ước lượng các trạng thái nhận thức. Cấu trúc thuật toán này vận hành như một mô hình phân tích độc lập, đồng thời thiết lập nền tảng lý thuyết cho việc tích hợp dữ liệu sinh trắc học ở các giai đoạn phát triển tiếp theo.

4.2.3 Cơ sở lựa chọn mô hình tính toán: Hệ cơ sở luật và mạng Bayes động

Nhằm giải quyết bài toán lượng hóa hành vi và điều hướng can thiệp, nghiên cứu thiết lập hai cấu trúc mô hình tính toán với các chức năng chuyên biệt: hệ cơ sở luật được ứng dụng để phân tích các mẫu hành vi thị giác, và mạng Bayes động được triển khai nhằm suy diễn trạng thái nhận thức của trẻ. Việc phân bổ nền tảng xử lý được điều chỉnh tương thích với

đặc điểm lâm sàng của từng phân nhóm. Cụ thể, hệ cơ sở luật được thiết kế cho nhóm trẻ rối loạn phổ tự kỉ mang đặc thù về hành vi tương tác, trong khi mạng Bayes động tập trung hỗ trợ nhóm có đặc thù về kỹ năng học tập (như kỹ năng đọc và tính toán). Căn cứ lựa chọn mô hình được xác lập qua các tiêu chí về biểu hiện hành vi và nhu cầu can thiệp như sau:

- Yêu cầu về tính minh bạch trong phân tích hành vi tương tác: Trong quá trình hỗ trợ đối tượng bằng các công cụ trực quan (như Picture Exchange Communication System - PECS, Treatment and Education of Autistic and related Communication-handicapped Children - TEACCH), chuyên gia cần xác định rõ cơ sở logic khi hệ thống đưa ra đề xuất điều chỉnh. Hệ cơ sở luật cung cấp các quy tắc suy diễn tường minh [157]. Cụ thể, khi dữ liệu ghi nhận tỷ lệ nhìn ra ngoài vùng mục tiêu vượt ngưỡng 80%, hệ thống xác định trạng thái phân tán chú ý và đề xuất kích hoạt tín hiệu âm thanh. Mặc dù các quy tắc hệ quả (IF-THEN) được thiết lập cố định, ngưỡng kích hoạt (θ) vẫn được tinh chỉnh theo phân phối cơ sở (μ, σ) của từng cá nhân nhằm đảm bảo yếu tố thích ứng.
- Yêu cầu về năng lực suy diễn đối với trạng thái nhận thức: Nhóm trẻ rối loạn phát triển đặc thù về kỹ năng học tập thường gặp rào cản trong quá trình xử lý thông tin nội tại (như giải mã âm vị hoặc tư duy toán học) thay vì các hành vi tương tác vật lý bên ngoài. Tiến trình này được biểu diễn dưới dạng một trạng thái ẩn, khó có thể đo lường trực tiếp thông qua các quy tắc tĩnh. Do đó, mô hình mạng Bayes động được áp dụng để tính toán xác suất mức độ làm chủ kiến thức ($P(L_t)$) dựa trên chuỗi dữ liệu tương tác liên tục của trẻ [153, 158].

4.3 Hệ thống can thiệp cá nhân hóa cho trẻ rối loạn phổ tự kỉ dựa trên cơ sở luật suy diễn và dữ liệu thị giác

4.3.1 Cấu trúc chu trình can thiệp năm giai đoạn tích hợp công nghệ theo dõi chuyển động mắt

Giải pháp can thiệp cho trẻ rối loạn phổ tự kỉ được cấu trúc theo một chu trình gồm năm giai đoạn. Điểm thay đổi chính của phương pháp này là

việc tích hợp công nghệ theo dõi chuyển động mắt vào quy trình giám sát nhằm hạn chế tính chủ quan của các phương pháp quan sát truyền thống. Cấu trúc tổng thể của chu trình được minh họa tại Hình 4.2.

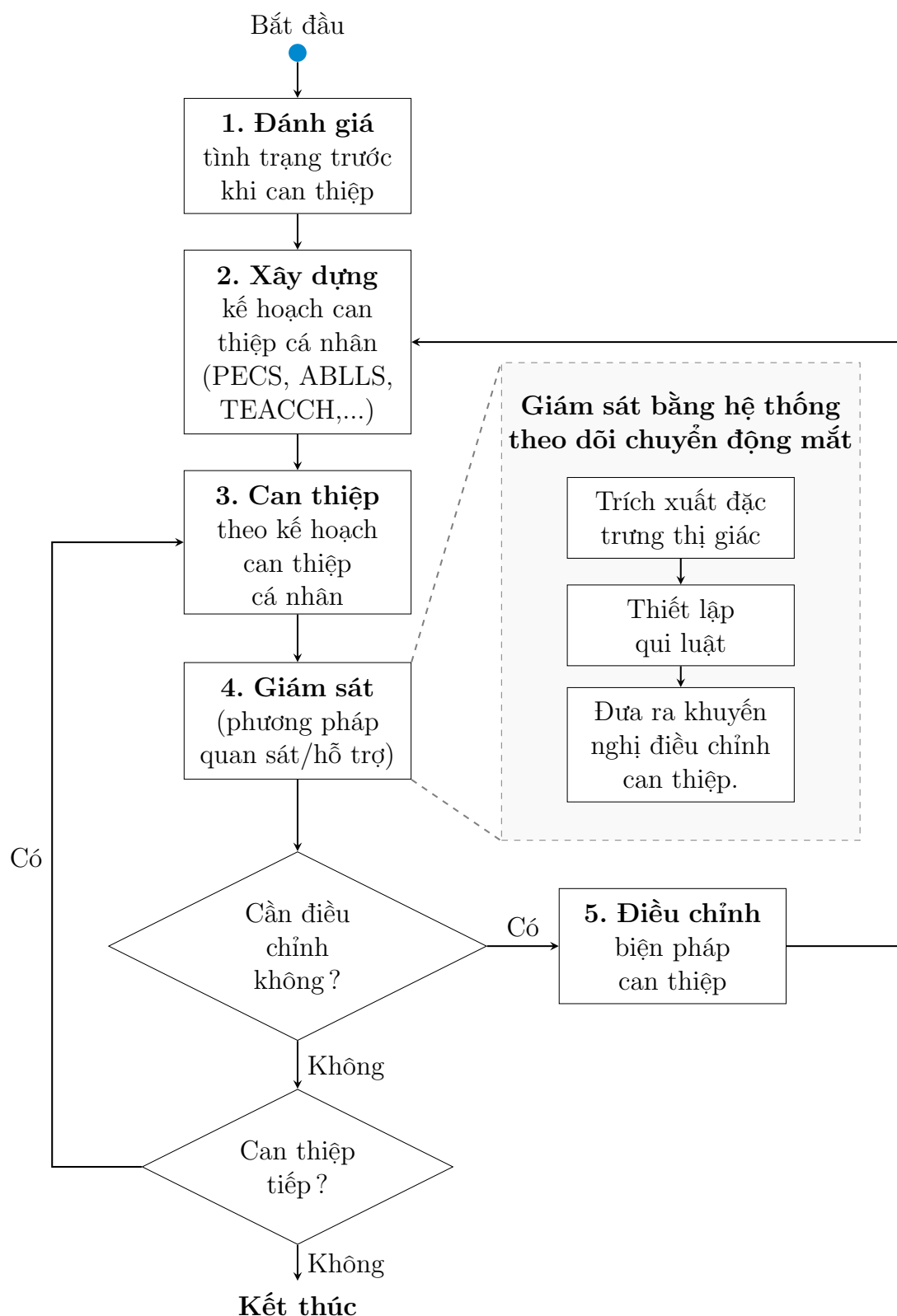
Các giai đoạn của tiến trình can thiệp bao gồm:

- Giai đoạn 1 và 2 (Đánh giá và lập kế hoạch): Thực hiện đánh giá năng lực nhận thức hiện tại của trẻ qua các công cụ tiêu chuẩn như PECS, Assessment of Basic Language and Learning Skills - Revised (ABLLS-R), TEACCH [159–161] để xây dựng kế hoạch can thiệp cá nhân hóa.
- Giai đoạn 3 (Thực thi can thiệp): Triển khai các nhiệm vụ tương tác (ví dụ: giao tiếp trao đổi hình ảnh, ghép câu) bám sát theo lộ trình đã định.
- Giai đoạn 4 (Giám sát tích hợp công nghệ): Ứng dụng thiết bị theo dõi chuyển động mắt để trích xuất luồng đặc trưng thị giác theo thời gian thực trong quá trình trẻ thực hiện nhiệm vụ. Chuỗi dữ liệu định lượng này được đối chiếu với hệ cơ sở luật [155] nhằm nhận diện trạng thái nhận thức, tạo cơ sở hỗ trợ quá trình ra quyết định lâm sàng.
- Giai đoạn 5 (Đề xuất hiệu chỉnh): Tổng hợp chuỗi phân tích từ hệ thống để đưa ra các khuyến nghị hỗ trợ. Việc quyết định phương án và thời điểm điều chỉnh được quyết định bởi chuyên gia, căn cứ vào bối cảnh thực tế. Cơ chế này thiết lập mô hình vòng lặp có sự tham gia của con người (human-in-the-loop), qua đó khẳng định vai trò quyết định của nhà chuyên môn.

4.3.2 Mô hình hóa hệ cơ sở luật suy diễn cá nhân hóa

Phân hệ giám sát và hệ thống khuyến nghị được xây dựng và phát triển dựa trên hệ cơ sở luật với cấu hình ngưỡng kích hoạt (θ) linh hoạt. Trước khi thực nghiệm, hệ thống thực hiện pha đánh giá cơ sở để thu thập dữ liệu và tính toán giá trị trung bình (μ) cùng độ lệch chuẩn (σ) về đặc trưng thị giác của từng trẻ. Các ngưỡng kích hoạt được thiết lập tự động theo công thức [162]:

$$\theta_{high} = \mu + k \cdot \sigma, \quad \theta_{low} = \mu - k \cdot \mu \quad (4.1)$$



Hình 4.2: Năm giai đoạn của chu trình can thiệp trẻ RLPTK tích hợp theo dõi chuyển động mắt

Hệ cơ sở luật tiếp nhận vectơ đặc trưng thị giác $V_t = [F_c, F_d, R_{in}, R_{out}, R_{AOI_{Distractor}}]$ tại thời điểm t làm dữ liệu đầu vào, trong đó ý nghĩa của các chỉ số được định nghĩa như sau:

- F_c : (*Fixation Count*): Tổng số lượng điểm cố định mắt trong một đơn vị thời gian cố định, phản ánh mật độ tìm kiếm thông tin trực quan.
- F_d : (*Fixation Duration*): Thời lượng trung bình của một điểm cố định mắt (tính bằng ms), thể hiện độ sâu xử lý nhận thức.
- R_{in} : Tỷ lệ thời gian ánh nhìn nằm trong phạm vi khung hình hiển thị nhiệm vụ học tập.
- R_{out} : Tỷ lệ thời gian ánh nhìn hướng ra ngoài khung hình, biểu thị xu hướng xao nhãng cơ học.
- $R_{AOI_{Distractor}}$: Tỷ lệ thời lượng điểm nhìn tập trung vào các vùng quan tâm gây nhiễu được thiết lập sẵn, phản ánh năng lực ức chế chú ý chọn lọc.

Dựa trên vectơ đặc trưng đầu vào, hệ thống thực hiện suy diễn để xác định trạng thái hành vi nhận thức S_t và đưa ra tín hiệu đề xuất hành động tương ứng A_t :

Luật 1: Nhận diện mắt tập trung

$$\begin{aligned} &\text{IF } (R_{out} > \theta_{out_{high}}) \wedge (F_d < \theta_{dur_{low}}) \\ &\text{THEN } S_t = \text{Mất tập trung} \\ &\Rightarrow A_t = \text{Đề xuất phát tín hiệu âm thanh điều hướng} \end{aligned} \quad (4.2)$$

Luật 2: Nhận diện suy giảm chú ý chọn lọc

$$\begin{aligned} &\text{IF } (R_{AOI_{Distractor}} > \theta_{AOI_{high}}) \wedge (F_c < \theta_{count_{low}}) \\ &\text{THEN } S_t = \text{Suy giảm chú ý chọn lọc} \\ &\Rightarrow A_t = \text{Đề xuất làm mờ vùng nhiễu trực quan} \end{aligned} \quad (4.3)$$

4.3.2.1 Cấu trúc hệ thống luật suy luận trạng thái hành vi

Sau quá trình trích xuất và định lượng đặc trưng thị giác, các quy tắc logic được thiết lập nhằm ánh xạ trực tiếp các chỉ số về thời lượng và tần suất điểm nhìn sang các trạng thái nhận thức tương ứng, diễn hình như

trạng thái phân tán hoặc suy giảm sự chú ý. Tiến trình này được triển khai thông qua ba bước chính:

- Bước 1: Số hóa luồng dữ liệu chuyển động mắt thành các biến định lượng thuộc vectơ V_t .
- Bước 2: Xác định danh mục các hành vi nhận thức dựa trên các biểu hiện lâm sàng của trẻ rối loạn phát triển.
- Bước 3: Áp dụng phương pháp Delphi qua các vòng khảo sát chuyên gia để xác định hệ số định hình ngưỡng tham chiếu k , hướng tới tính đồng thuận khoa học.

Trên cơ sở các đặc trưng thị giác đã được định lượng và các ngưỡng tham chiếu thiết lập, hệ thống tiến hành thiết kế các luật chuyển đổi sang các hành vi nhận thức thông qua các phép toán logic có điều kiện. Ví dụ:

$$\mathbf{IF} (F_{dDistractor} > \theta_{distractor_{high}}) \mathbf{THEN} S_t = \text{Phân tán sự chú ý} \quad (4.4)$$

Trong đó:

- $F_{dDistractor}$: Thời lượng điểm cố định trung bình của trẻ trên các vùng chứa yếu tố gây nhiễu trực quan.
- $\theta_{distractor_{high}}$: Ngưỡng giới hạn trên của thời lượng nhìn vật nhiễu, được cá nhân hóa theo từng đối tượng. Đây là mức thời gian tối đa một đối tượng cụ thể có thể giới hạn việc duy trì ánh nhìn vào các yếu tố gây xao nhãng trên khung hình tương tác trước khi hệ thống xác định trạng thái mất tập trung và đưa ra khuyến nghị can thiệp.

Cuối cùng, hệ thống luật được hiệu chỉnh thông qua quy trình đánh giá phối hợp với hội đồng chuyên gia về giáo dục đặc biệt nhằm hỗ trợ tăng cường độ chính xác và giảm thiểu tỷ lệ dương tính giả.

4.3.2.2 Mô hình ánh xạ đề xuất hiệu chỉnh và tối ưu hóa can thiệp

Tiến trình này vận hành như một chức năng hỗ trợ ra quyết định để xây dựng các phương án can thiệp. Về mặt toán học, tiến trình được biểu diễn bằng hàm ánh xạ $R = f(S_t)$, chuyển đổi trạng thái nhận thức S_t thành danh mục đề xuất cá nhân hóa R . Hệ thống không tự động thực thi các hành

động này nhằm hạn chế các kích thích không phù hợp cho trẻ. Mọi khuyến nghị đều được chuyên gia hoặc giáo viên trực tiếp thẩm định để đảm bảo phù hợp với bối cảnh lâm sàng và thực tế của trẻ.

Việc khai thác các đặc trưng thị giác cung cấp nguồn thông tin khách quan để hỗ trợ quy trình đánh giá và xác định phương án hỗ trợ. Cơ chế tương tác giữa hệ thống và chuyên gia được thể hiện qua hai trường hợp:

- Đối với dấu hiệu suy giảm chú ý xã hội: Tại thời điểm ghi nhận, hệ thống gửi tín hiệu cảnh báo đến giáo viên và nhà can thiệp để lưu ý trạng thái hiện thời của trẻ. Kết thúc phiên học, hệ thống tự động tổng hợp toàn bộ các thời điểm xảy ra hành vi bất thường, từ đó đưa ra đề xuất tổng thể về việc điều chỉnh phương pháp can thiệp hoặc bổ sung các nhiệm vụ tăng cường tương tác mắt cho phiên can thiệp kế tiếp.
- Đối với trạng thái lệch hướng tập trung: Khi phát hiện ánh nhìn của trẻ rời xa vùng mục tiêu, một thông báo tức thời được chuyển đến người can thiệp để theo dõi ngữ cảnh lâm sàng. Sau đó, dữ liệu của cả phiên đo được phân tích để phân định bản chất của sự xao nhãng (nhất thời hay hệ thống), cung cấp luận cứ thực chứng giúp chuyên gia cân nhắc và đưa ra quyết định thay đổi cấu trúc học liệu phù hợp trước khi bắt đầu phiên can thiệp tiếp theo.

Các kết quả về đặc trưng thị giác, trạng thái hành vi và đề xuất can thiệp tương ứng được hệ thống hóa tại Bảng 4.2.

4.3.3 Phương pháp đánh giá hiệu quả can thiệp thông qua thiết kế thực nghiệm trường hợp đơn và thống kê Tau-U

Trong quá trình can thiệp, luận án đánh giá sự thay đổi của hành vi thị giác thông qua các nhiệm vụ tương tác có cấu trúc. Thay vì so sánh giữa các nhóm đối tượng, luận án áp dụng thiết kế. Phương pháp này đo lường lặp lại hành vi của từng cá nhân qua các giai đoạn liên tiếp, xem mỗi trẻ là một đơn vị phân tích độc lập. Quy trình thực nghiệm gồm hai pha: pha cơ sở (A) nhằm ghi nhận hành vi khi chưa có tác động, và pha can thiệp (B) để đánh giá hành vi sau khi áp dụng biện pháp hỗ trợ.

Hiệu quả can thiệp được lượng hóa thông qua kiểm định Tau-U. Đây là phương pháp thống kê phi tham số dành cho thiết kế SCED, dựa trên việc

Bảng 4.2: Các chỉ số, cơ chế suy luận logic và diễn giải hành vi trong hệ thống hỗ trợ can thiệp cho trẻ rối loạn phát triển

Khía cạnh hành vi	Chỉ số kỹ thuật đại diện	Điều kiện toán học	Diễn giải lâm sàng và Khuyến nghị hỗ trợ (Chuyên gia quyết định)
Sự tham gia tổng thể	Tỷ lệ phân bổ ánh nhìn (R_{trong}, R_{ngoai})	$R_{trong} > \theta_{trong_cao}$	Tăng cường năng lực chú ý: trẻ tương tác tốt với khung hình nhiệm vụ. Khuyến nghị giữ nguyên tiến độ hoặc tăng mức độ củng cố hành vi.
		$R_{trong} \in [\theta_{thap}, \theta_{cao}]$	Chậm trễ chú ý: Xu hướng mệt mỏi hoặc suy giảm động lực học tập. Khuyến nghị xem xét rút ngắn thời lượng phiên học và kiểm tra tiếng ồn môi trường.
		$R_{ngoai} > \theta_{ngoai_cao}$	Xao nhãng chú ý: Ánh nhìn rời khỏi không gian hiển thị bài học. Khuyến nghị phát tín hiệu âm thanh điều hướng hoặc thực hiện can thiệp vật lý trực tiếp.
Chiến lược xử lý	Mật độ và độ sâu điểm cố định (F_c, F_d)	$F_c > \theta_{mat_do_cao} \wedge F_d < \theta_{do_sau_thap}$	Hiện tượng quét quá mức: Tình trạng quá tải nhận thức hoặc bối rối trước học liệu. Khuyến nghị phân rã cấu trúc nhiệm vụ.
		$F_c < \theta_{mat_do_thap} \wedge F_d > \theta_{do_sau_cao}$	Xử lý nhận thức sâu: Khả năng chú ý ổn định và tiếp thu tốt bản chất nhiệm vụ. Khuyến nghị tăng dần độ phức tạp của nhiệm vụ.
Chú ý có chọn lọc	Năng lực ức chế tác nhân nhiễu ($R_{AOI_Nhiều}$)	$R_{AOI_Nhiều} < \theta_{nhiều_thap}$	Ức chế nhiễu tốt: Hiệu suất tập trung chọn lọc cao. Khuyến nghị duy trì cấu trúc giao diện hiện tại.
		$R_{AOI_Nhiều} > \theta_{nhiều_cao}$	Suy giảm chú ý chọn lọc: Chú ý thị giác bị thu hút bởi các chi tiết không liên quan. Khuyến nghị làm mờ vùng nhiễu trực quan trên giao diện phần mềm.
Độ thành thạo kỹ năng	Thời gian phản hồi (T_{res}) và Độ chính xác (ACC)	$ACC > \theta_{cx_cao} \wedge T_{res} < \theta_{tg_thap}$	Tự động hóa hành vi: Kỹ năng nhận thức đạt độ phản xạ trôi chảy. Khuyến nghị chuyển sang các dạng nhiệm vụ khái quát hóa năng lực.
		$ACC > \theta_{cx_cao} \wedge T_{res} > \theta_{tg_cao}$	Phản hồi chính xác chậm: Trạng thái hoàn thành nhiệm vụ còn do dự. Khuyến nghị bổ sung các phiên rèn luyện tương đồng để tăng cường phản xạ chính xác.

tính toán tỷ lệ phân phối dữ liệu không chồng lấp giữa pha A và pha B, đồng thời kiểm soát xu hướng của đường cơ sở [142]. Căn cứ theo hệ thống phân loại trong giáo dục đặc biệt, giá trị Tau-U từ 0,66 đến 0,92 thể hiện mức độ tác động đáng kể, và từ 0,93 trở lên phản ánh mức độ tác động rất lớn [163].

4.3.3.1 Đặc tả cấu trúc và ý nghĩa của dữ liệu thô đầu vào

Dữ liệu đầu vào được thu thập độc lập từ ba trẻ (P1, P2, P3), tương ứng với ba nhiệm vụ đo lường hành vi thị giác:

- Trẻ P1 (Nhiệm vụ quan sát tự do): trẻ quan sát khung hình chứa các đối tượng quen thuộc không có mục tiêu chỉ định. Tham số ghi nhận là tổng thời lượng điểm cố định trong một phiên.
- Trẻ P2 (Nhiệm vụ quan sát AOI mục tiêu): trẻ quan sát đối tượng mục tiêu được chỉ định giữa các yếu tố nhiễu thị giác. Tham số ghi nhận là tỷ lệ thời gian cố định ánh nhìn tại vùng mục tiêu.
- Trẻ P3 (Nhiệm vụ ghép câu): trẻ quan sát các đối tượng hiển thị theo bố cục phân lớp thành phần ngữ pháp, sau đó tiến hành ghép nối để thành lập câu hoàn chỉnh. Tham số ghi nhận là chuỗi thứ tự các AOI được kích hoạt.

Hệ thống ghi nhận từng điểm nhìn cố định qua các tham số: mã vùng quan tâm, thời điểm bắt đầu, thời điểm kết thúc và thời lượng tính bằng mili giây (ms). Sự kết hợp giữa vị trí vùng quan tâm và thời lượng điểm nhìn là cơ sở để hệ thống xác định mức độ tập trung và chiến lược nhận thức của trẻ.

4.3.3.2 Quy trình trích xuất và thiết lập thang đo nhận thức

Hệ thống xử lý các bản ghi luồng dữ liệu điểm nhìn nhằm trích xuất các sự kiện thị giác. Thời lượng của từng điểm nhìn cố định được tính toán dựa trên công thức:

$$\text{Duration}_{ms} = \text{End}_{ms} - \text{Start}_{ms} \quad (4.5)$$

Tiếp đó, các thuật toán lọc được triển khai để loại bỏ tín hiệu nhiễu phát sinh từ chuyển động chớp mắt và các sai số kỹ thuật của thiết bị.

Kết thúc giai đoạn tiền xử lý, chuỗi dữ liệu được chuyển thành các tham số đặc trưng tương ứng với từng nhiệm vụ thực nghiệm, bao gồm 80 phiên

đo trên mỗi trẻ. Cụ thể, đối với trẻ P1, hệ thống tổng hợp chỉ số thời lượng từ 1.470 điểm nhìn cố định. Đối với trẻ P2 và P3, dữ liệu được quy đổi sang thang đo tỷ lệ phần trăm nhằm cung cấp cơ sở đánh giá mức độ chính xác trong quá trình duy trì sự chú ý vào vùng quan tâm mục tiêu.

4.3.3.3 Mô hình giả thuyết và phương pháp kiểm định Tau-U

Dữ liệu thực nghiệm được hệ thống ghi nhận thông qua chuỗi tương tác kéo dài 80 phiên đo liên tục. Dựa trên phương pháp SCED, chuỗi thực nghiệm này được phân chia theo một cấu trúc cố định bao gồm hai giai đoạn: pha cơ sở A (từ phiên 1 đến phiên 25) và pha can thiệp B (từ phiên 26 đến phiên 80).

Để đánh giá hiệu quả của hệ thống, luận án thiết lập giả thuyết vô hiệu (H_0) với giả định rằng quá trình can thiệp không tạo ra hiệu ứng làm thay đổi hành vi của trẻ. Về mặt thống kê, điều này đồng nghĩa với việc không tồn tại sự khác biệt có ý nghĩa về phân phối dữ liệu giữa pha cơ sở A và pha can thiệp B. Nhằm kiểm chứng giả thuyết này, toàn bộ dữ liệu được xử lý thông qua phương pháp kiểm định phi tham số Tau-U [142] để tính toán chỉ số Z và giá trị p (kiểm định hai đuôi). Cơ sở định lượng này cho phép xác định mức độ tác động của quá trình can thiệp để đưa ra quyết định bác bỏ hoặc chấp nhận H_0 . Đồng thời, phương pháp phân tích biểu đồ trực quan được triển khai song song để đối chiếu xu hướng, mức độ dịch chuyển và tính ổn định của chuỗi dữ liệu khi trẻ chuyển tiếp giữa các pha thực nghiệm.

4.3.3.4 Phân tích kết quả thực nghiệm trường hợp đơn

Bảng 4.3 tổng hợp các chỉ số thống kê về hành vi thị giác của ba trẻ thu được từ quá trình thực nghiệm.

Phân tích thống kê và phân tích thị giác cho thấy các đặc trưng về hiệu quả can thiệp sau:

- Đối với trẻ P1 (Nhiệm vụ 1 – Quan sát tự do):
 - Hành vi thị giác tại pha cơ sở: Tổng thời lượng quan sát biến thiên trong khoảng từ 1.000 *ms* đến 2.000 *ms*.
 - Hành vi thị giác tại pha can thiệp và kết quả thống kê: Tổng thời lượng quan sát có xu hướng tăng tuyến tính, nằm trong khoảng từ

Bảng 4.3: Chỉ số hành vi thị giác và phân tích kiểm định tích Tau-U (SCED 80 phiên)

Trẻ	Chỉ số đo	Trung bình pha A	Trung bình pha B	Tau-U	Z	p-value
P1 (Nhiệm vụ 1 - Nhìn tự do)	Tổng thời lượng (ms)	1567.96	4089.07	0.8422	6.01	< 0.001
P2 (Nhiệm vụ 2 - Quan sát AOI mục tiêu)	Tỉ lệ nhìn đích (%)	20.10%	42.58%	0.9200	6.57	< 0.001
P3 (Nhiệm vụ 3 - Thành lập câu)	Tỉ lệ thời gian nhìn đúng hướng tiến triển câu (%)	47.88%	58.95%	0.9404	6.71	< 0.001

6.000 *ms* đến 9.000 *ms*. Kiểm định Tau-U đạt 0,8422 ($p < 0,001$), thể hiện sự thay đổi có hệ thống với phân phối thời lượng điểm nhìn tập trung rõ ràng hơn ở khoảng từ 200 *ms* đến 800 *ms*.

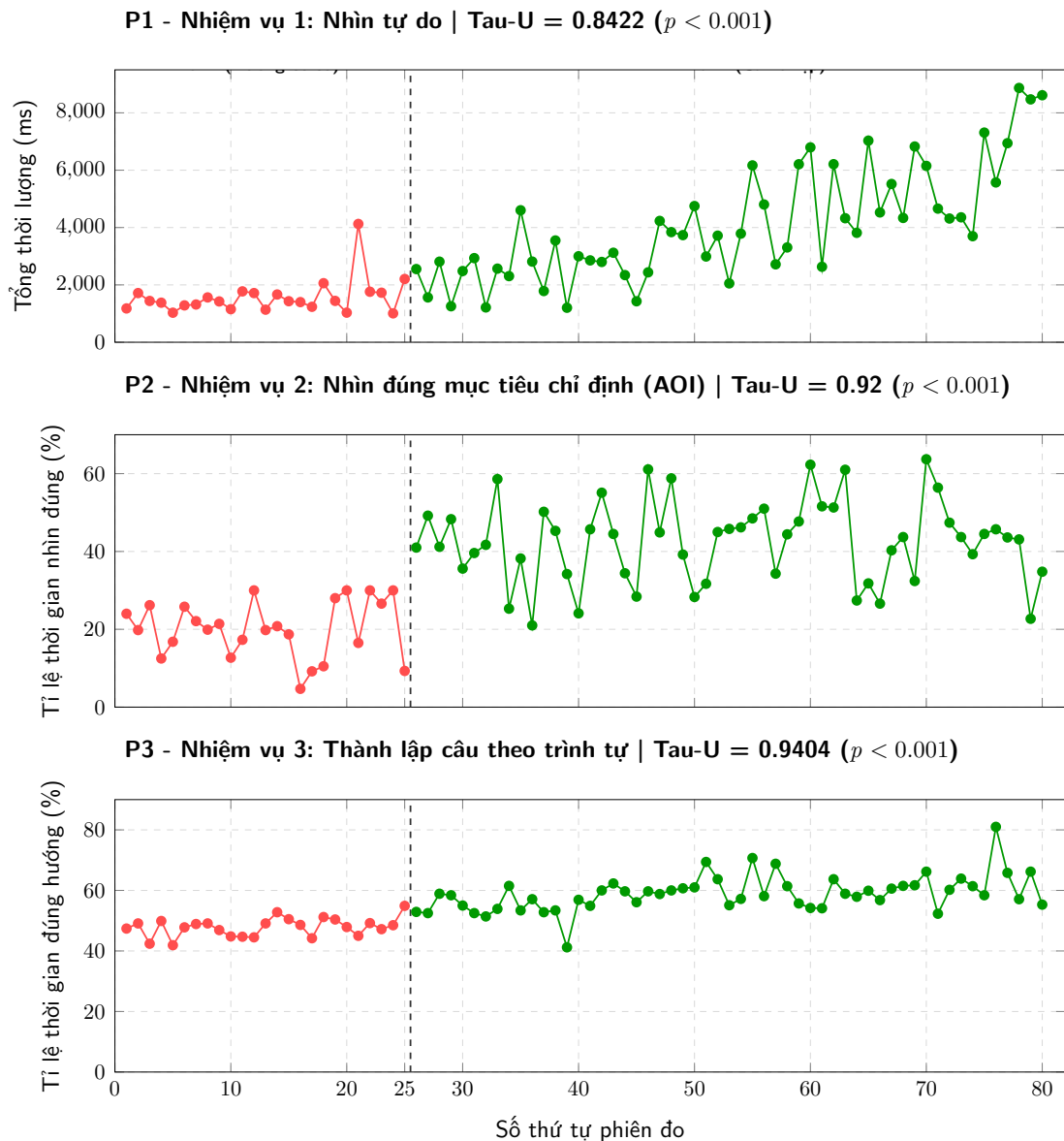
– Đối với trẻ P2 (Nhiệm vụ 2 – Quan sát AOI mục tiêu):

- Hành vi thị giác tại pha cơ sở: Tỷ lệ thời gian duy trì ánh nhìn vào mục tiêu dao động ở mức từ 5% đến 30% với thời lượng điểm nhìn trung bình ở mức 107,7 *ms*.
- Hành vi thị giác tại pha can thiệp và kết quả thống kê: Tỷ lệ thời gian nhìn mục tiêu tăng lên mức 40% ngay từ phiên đầu tiên của pha B so với mức trung bình 20,10% của pha A. Kiểm định Tau-U đạt 0,9200 ($p < 0,001$), đồng thời thời lượng điểm nhìn trung bình dịch chuyển lên mức 140,1 *ms*, hỗ trợ luận cứ về một tiến trình xử lý thông tin sâu hơn.

– Đối với trẻ P3 (Nhiệm vụ 3 - Ghép câu theo trình tự):

- Hành vi thị giác tại pha cơ sở: Tỷ lệ thời gian nhìn theo đúng hướng tiến triển cấu trúc câu (từ chủ ngữ đến trợ động từ, động từ và đối tượng) dao động trong khoảng từ 40% đến 55%, trung bình 47,88%.
- Hành vi thị giác tại pha can thiệp và kết quả thống kê: Tỷ lệ thời gian nhìn đúng hướng tiến triển câu tăng và duy trì trong khoảng từ 50% đến 80%, trung bình 58,95%. Kiểm định Tau-U đạt 0,9404 ($p < 0,001$), thể hiện khả năng tổ chức trình tự quan sát theo cấu trúc ngôn ngữ ổn định hơn sau can thiệp.

Nhìn chung, sự dịch chuyển xu hướng tại thời điểm chuyển pha cùng tỷ lệ chồng lấp dữ liệu thấp giữa hai giai đoạn cho thấy các biện pháp can



Hình 4.3: Biểu đồ phân tích thị giác chuỗi thời gian (SCED) của ba trẻ.

thiệp hành vi thị giác hướng tới sự chuyển dịch tích cực, mặc dù có sự khác biệt về cấu trúc của từng nhiệm vụ tương tác. Những thay đổi trong xu hướng cũng như mức độ phân bố dữ liệu qua các giai đoạn được trực quan hóa tại Hình 4.3. Đường nét đứt tại phiên thứ 25 đánh dấu ranh giới giữa pha đường cơ sở (A) và pha can thiệp (B). Cả ba trẻ đều thể hiện mức điểm trong pha B cao hơn so với pha A, với giá trị Tau-U dao động từ 0,8422 đến 0,9404, đều đạt ý nghĩa thống kê $p < 0,001$ ở cả ba trường hợp, giả thuyết vô hiệu H_0 bị bác bỏ, cho thấy quá trình can thiệp tạo ra sự thay đổi có hệ thống và không ngẫu nhiên trong hành vi thị giác của từng trẻ, thể hiện

qua sự dịch chuyển đáng kể giữa pha cơ sở A và pha can thiệp B trên cả ba nhiệm vụ thực nghiệm.

4.4 Các nền tảng số hỗ trợ can thiệp trẻ rối loạn tính toán và trẻ mắc chứng khó đọc

Bên cạnh hệ cơ sở luật dành cho trẻ RLPTK, luận án phát triển hai nền tảng số để hỗ trợ nhóm trẻ mắc chứng khó đọc và rối loạn tính toán. Ở giai đoạn hiện tại, các nền tảng này chưa tích hợp thiết bị theo dõi chuyển động mắt. Trọng tâm của hệ thống là điều chỉnh cấu hình tương tác dựa trên thuyết tải nhận thức [45] nhằm giảm độ phức tạp của giao diện và hạn chế ảnh hưởng các biến số gây nhiễu chú ý. Quá trình thao tác của trẻ trên các giao diện này được hệ thống ghi nhận thành các bộ dữ liệu hành vi có cấu trúc bao gồm thời gian phản hồi, tỷ lệ lỗi và chuỗi thao tác. Chuỗi dữ liệu này đóng vai trò làm dữ liệu thực chứng, tạo tiền đề cho việc xây dựng và huấn luyện các mô hình học máy ước lượng trạng thái nhận thức ở các giai đoạn tiếp theo.

4.4.1 Nền tảng web EzMath hỗ trợ trẻ rối loạn tính toán

Đối với nhóm trẻ RLTT, các rào cản nhận thức chủ yếu bao gồm sự khó khăn trong tư duy logic, sự khiếm khuyết về kỹ năng tổ chức không gian thị giác và giới hạn ghi nhớ của trí nhớ làm việc [164]. Dựa trên các đặc điểm lâm sàng này, nền tảng EzMath tái cấu trúc không gian hiển thị theo hai nguyên tắc thiết kế trọng tâm:

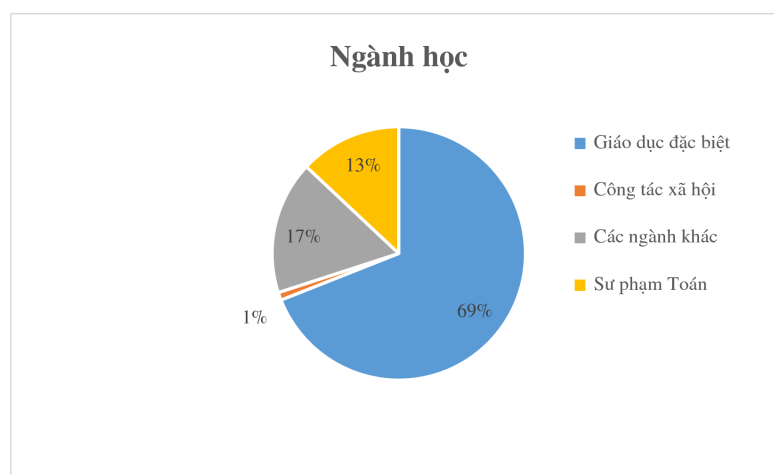
- Giảm nhiễu thông tin và tăng cường phân lập thị giác: Các chữ số được bố trí với độ giãn cách rộng nhằm khắc phục hiện tượng chồng lấn thị giác, một rào cản sinh lý khiến trẻ gặp khó khăn trong việc phân biệt các ký tự khi chúng nằm quá sát nhau [165]. Phối hợp với quy tắc này, hệ thống áp dụng kỹ thuật phân mã màu nền để định vị tách biệt các hàng giá trị (hàng chục và hàng đơn vị), qua đó hỗ trợ tiến trình phân rã và nhận diện cấu trúc số học của trẻ.
- Giảm thiểu rào cản thao tác vận động: Nhận thấy sự hạn chế về kỹ năng vận động tinh có thể tạo ra áp lực ngoại lai và làm phân tán sự chú ý của trẻ trong khâu nhập liệu [55], hệ thống thiết kế các phương

thức tương tác thay thế cho bàn phím tiêu chuẩn. Thao tác trả lời được chuyển đổi thành các hành vi tối giản thông qua bộ nút điều chỉnh giá trị, thao tác kéo thả đối tượng hoặc lệnh điều khiển bằng giọng nói.

Đánh giá thực nghiệm nền tảng EzMath

Để đánh giá hiệu quả của nền tảng EzMath, luận án đã thực hiện khảo sát nhằm thu thập ý kiến phản hồi của những người có chuyên môn hoặc có liên quan đến lĩnh vực giáo dục đặc biệt và hỗ trợ trẻ có khó khăn trong học tập môn Toán.

Đối tượng tham gia gồm 73 giáo viên, nhà nghiên cứu, học viên và sinh viên thuộc lĩnh vực giáo dục đặc biệt, phương pháp giảng dạy Toán, giáo dục người khuyết tật và các lĩnh vực liên quan tại các cơ sở giáo dục trên cả nước. Cơ cấu nghề nghiệp và ngành học hoặc lĩnh vực chuyên môn của nhóm tham gia được trình bày tương ứng tại Hình 4.4.



Hình 4.4: Ngành học hoặc lĩnh vực chuyên môn của người tham gia khảo sát

Dựa trên phản hồi của người tham gia, hệ thống đã được đánh giá tập trung vào giao diện người dùng và khả năng hỗ trợ trẻ em có khó khăn trong học tập môn Toán.

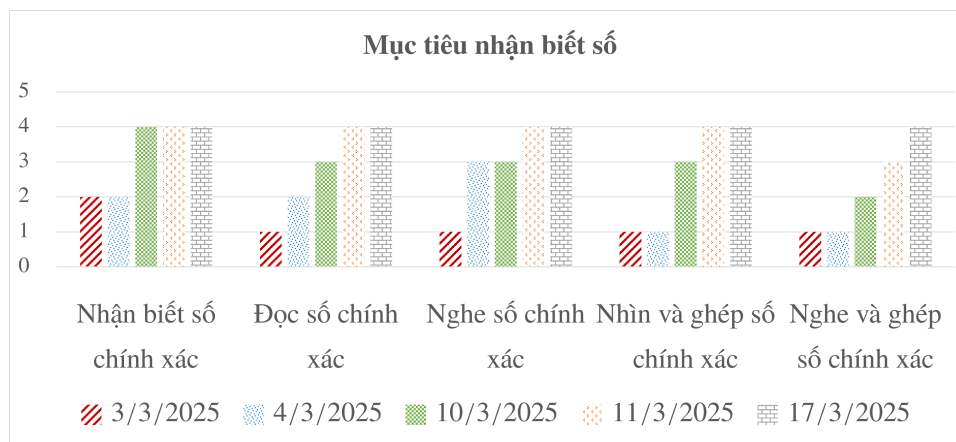
Giao diện người dùng được đánh giá dựa trên khả năng cung cấp thông tin và hỗ trợ tương tác, với các tiêu chí như: tính dễ sử dụng, tính rõ ràng, tính thẩm mỹ và mức độ phù hợp với đối tượng người dùng mục tiêu. Ở nhóm tiêu chí này, hệ thống nhận được mức đánh giá cao, với điểm số cho các tiêu chí về giao diện dao động từ 4,1 đến 4,9 trên thang điểm 5.

Khả năng hỗ trợ của hệ thống được đánh giá dựa trên năng lực cung cấp các hình thức hỗ trợ học tập đa dạng và hấp dẫn, bao gồm: thể hiện phù hợp đặc điểm của học sinh, thu hút sự chú ý, phù hợp với sở thích của trẻ và khuyến khích trẻ chủ động khám phá trong quá trình học tập. Mức độ hài lòng của người dùng đối với các tính năng hỗ trợ do hệ thống cung cấp đạt mức rất cao, với điểm số từ 4,6 đến 4,8.

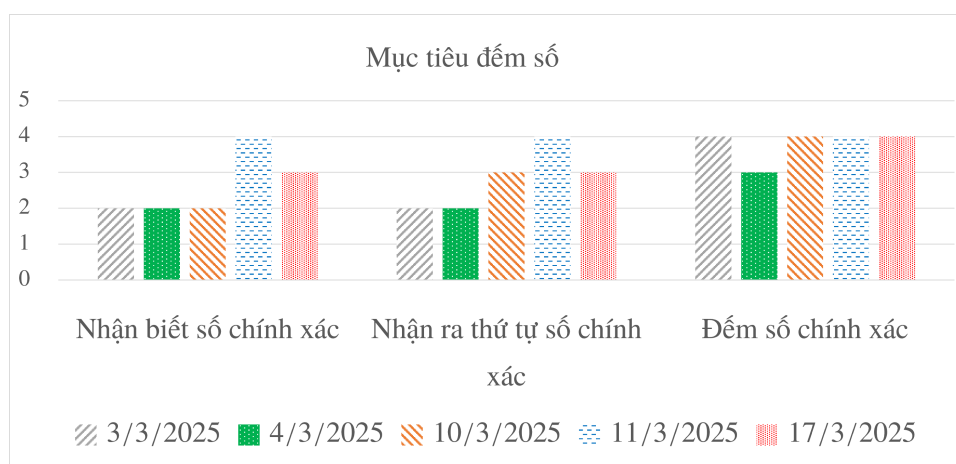
Nhìn chung, mức độ phù hợp của hệ thống đối với trẻ em có khó khăn trong học tập đạt điểm trung bình 4,42. Kết quả này xếp hệ thống vào mức "Phù hợp" đến "Rất phù hợp" (được xác định là từ 4,1 đến 5,0 điểm), đồng thời khẳng định ý nghĩa sư phạm và giá trị ứng dụng của hệ thống đối với nhóm người dùng mục tiêu

Để đánh giá hiệu quả của nền tảng EzMath trong việc cải thiện kỹ năng toán học, luận án tiến hành can thiệp cá nhân cho một học sinh lớp 1 RLTT, do một giáo viên giáo dục đặc biệt trực tiếp hướng dẫn thông qua thiết bị máy tính. Chương trình can thiệp được xây dựng theo hướng cá nhân hóa, điều chỉnh nội dung theo nhu cầu, đặc điểm và sở thích riêng của trẻ với tần suất 2 buổi/tuần, mỗi buổi kéo dài từ 30 đến 60 phút trong suốt học kỳ. Kết quả được ghi nhận tại 5 mốc thời gian từ ngày 3/3/2025 đến 17/3/2025 và phản ánh qua bốn nhóm mục tiêu như sau:

- Mục tiêu nhận biết số (Hình 4.5): Ở giai đoạn đầu can thiệp (3/3/2025), hầu hết các tiêu chí chỉ đạt mức 1–2/5. Cụ thể, *nhận biết số chính xác* đạt 2/5, trong khi các tiêu chí còn lại (*đọc số, nghe số, nhìn và ghép số, nghe và ghép số*) đều ở mức 1/5. Từ buổi 10/3/2025, phần lớn các tiêu chí tăng lên mức 3–4/5. Đến ngày 17/3/2025, cả năm tiêu chí đều đạt mức 4/5, cho thấy các kỹ năng nhận biết số cơ bản của trẻ đã được cải thiện và mức độ hỗ trợ giảm xuống.
- Mục tiêu đếm số (Hình 4.6): Tiêu chí *đếm số chính xác* có xuất phát điểm cao nhất (4/5 ngay từ 3/3/2025) và duy trì ổn định ở mức 4/5 trong suốt quá trình can thiệp, chỉ giảm nhẹ xuống 3/5 vào ngày 4/3/2025 rồi trở lại 4/5. Tiêu chí *nhận biết số chính xác* tăng từ 2/5 lên 4/5 từ ngày 11/3/2025 nhưng giảm nhẹ về 3/5 vào 17/3/2025. Tiêu chí *nhận ra thứ tự số chính xác* tiến triển chậm hơn, tăng từ 2/5 lên 3–4/5 từ ngày 10/3/2025 và ổn định ở mức 3/5 vào cuối giai đoạn. Kết quả này cho thấy "kỹ năng đếm" của trẻ được duy trì và củng cố theo thời gian khi sử dụng nền tảng web.

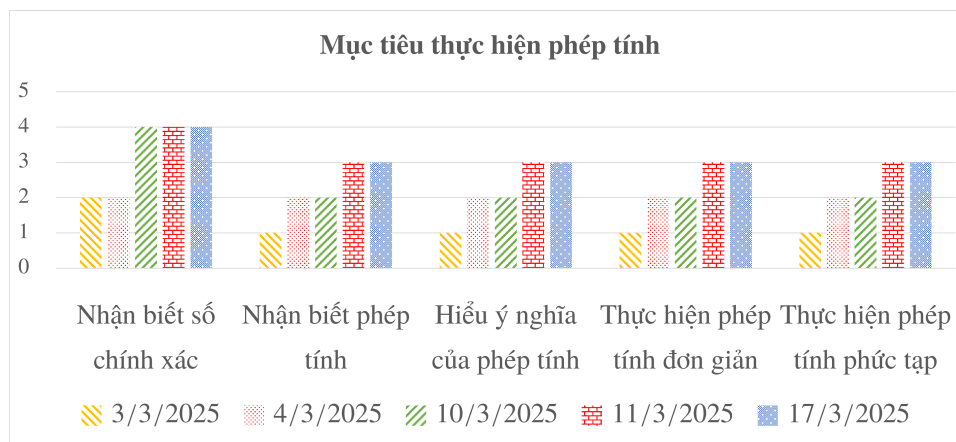


Hình 4.5: Số liệu thống kê mục tiêu nhận biết và xử lý số



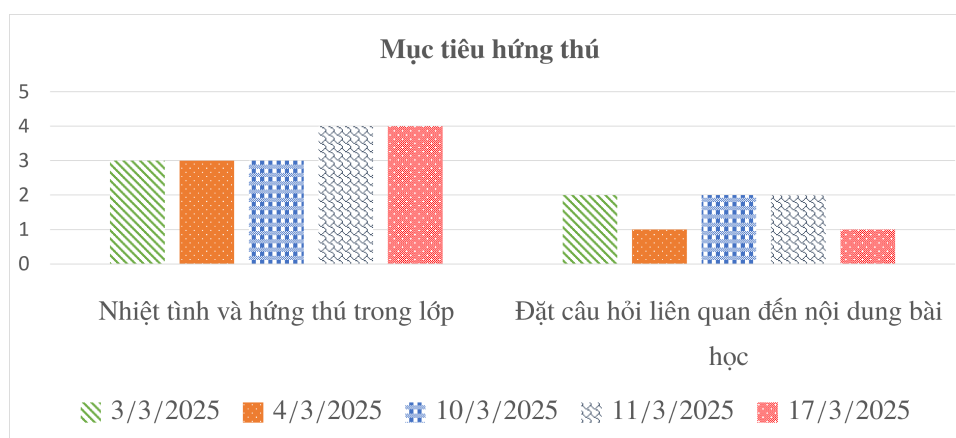
Hình 4.6: Số liệu thống kê mục tiêu đếm số

- Mục tiêu thực hiện phép tính (Hình 4.7): Đây là nhóm mục tiêu mà trẻ có xuất phát điểm thấp nhất. Ở buổi đầu (03/03/2025), các tiêu chí *nhận biết phép tính*, *hiểu ý nghĩa của phép tính*, *thực hiện phép tính đơn giản* và *thực hiện phép tính phức tạp* đều ở mức 1/5, chỉ riêng *nhận biết số chính xác* đạt 2/5. Đến buổi 4/3/2025, tất cả các tiêu chí tăng lên mức 2/5. Từ ngày 10/3/2025, tiêu chí *nhận biết số chính xác* đạt 4/5 và duy trì đến cuối giai đoạn. Các tiêu chí còn lại tăng lên mức 3/5 từ ngày 11/3/2025 và giữ ổn định đến 17/3/2025. Trước khi sử dụng ứng dụng, trẻ thường xuyên dùng hầu hết các mức hỗ trợ, đặc biệt hỗ trợ đọc được sử dụng trong tất cả các yêu cầu. Sau khi hệ thống bổ sung tính năng minh họa động cho phép cho thấy trẻ có xu hướng sử dụng tính năng này thường xuyên hơn.
- Mục tiêu hứng thú (Hình 4.8): Tiêu chí *nhật tình và hứng thú trong lớp* duy trì ở mức 3/5 trong ba buổi đầu (3/3 đến 10/3/2025) và tăng



Hình 4.7: Số liệu thống kê mục tiêu thực hiện phép tính

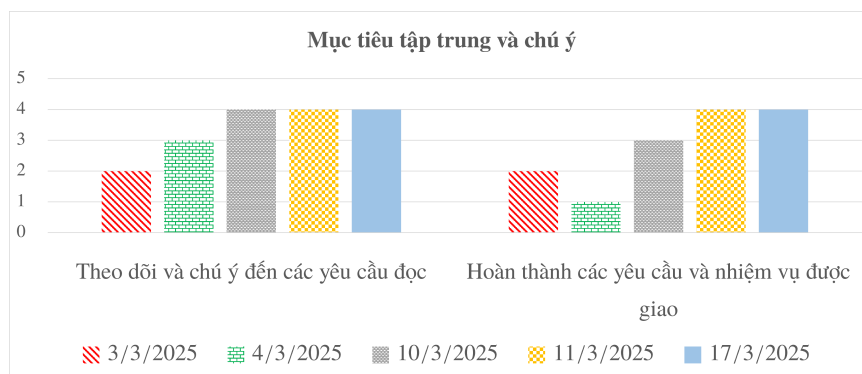
lên 4/5 từ ngày 11/3/2025, cho thấy mức độ hứng thú của trẻ trong các buổi học trên nền tảng có xu hướng tăng. Tiêu chí *đặt câu hỏi liên quan đến nội dung bài học* dao động trong khoảng 1–2/5 suốt quá trình can thiệp, mức độ chủ động đặt câu hỏi của trẻ chưa cải thiện rõ rệt và cần được hỗ trợ trong các giai đoạn tiếp theo.



Hình 4.8: Số liệu thống kê mục tiêu hứng thú

- Mục tiêu tập trung và chú ý (Hình 4.9): Tiêu chí *theo dõi và chú ý đến các yêu cầu đọc* cải thiện rõ rệt, tăng từ 2/5 (3/3/2025) lên 3/5 (4/3/2025) và đạt mức 4/5 từ ngày 10/3/2025, duy trì ổn định đến cuối giai đoạn. Tiêu chí *hoàn thành các yêu cầu và nhiệm vụ được giao* có phần dao động ở giai đoạn đầu (giảm từ 2/5 xuống 1/5 vào 4/3/2025), nhưng tăng mạnh lên 3/5 vào 10/3/2025 và đạt mức 4/5 ổn định từ ngày 11/3/2025. So với phương pháp truyền thống sử dụng nhiệm vụ trên giấy với hình thức cố định khó tùy chỉnh, trẻ thể hiện sự hứng thú và gia tăng mức độ tập trung khi học trên nền tảng web

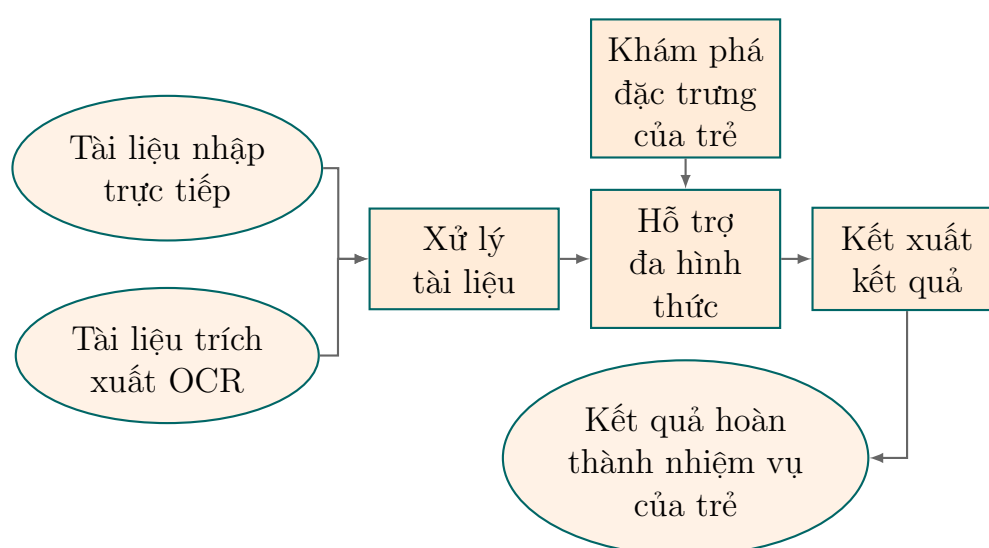
EzMath. Bên cạnh đó, việc bổ sung hình thức trả lời bằng giọng nói giúp trẻ hoàn thành yêu cầu minh chứng cho sự gia tăng mức độ tương tác của trẻ đối với nền tảng.



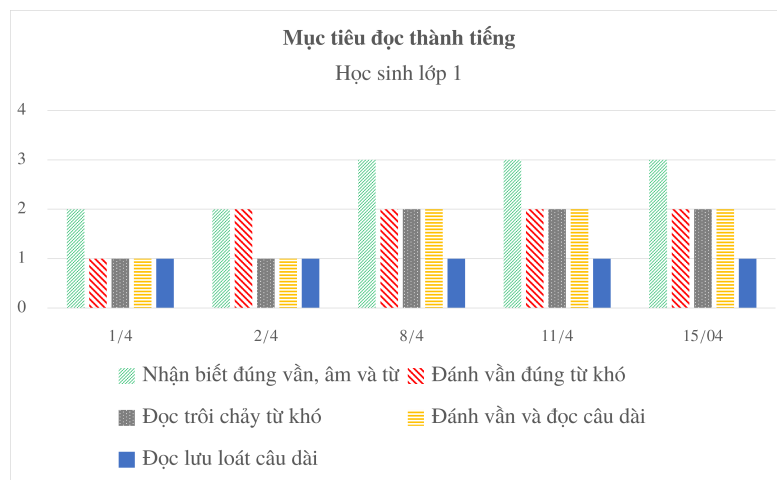
Hình 4.9: Số liệu thống kê mục tiêu tập trung và chú ý

4.4.2 Nền tảng di động EzRead hỗ trợ trẻ mắc chứng khó đọc

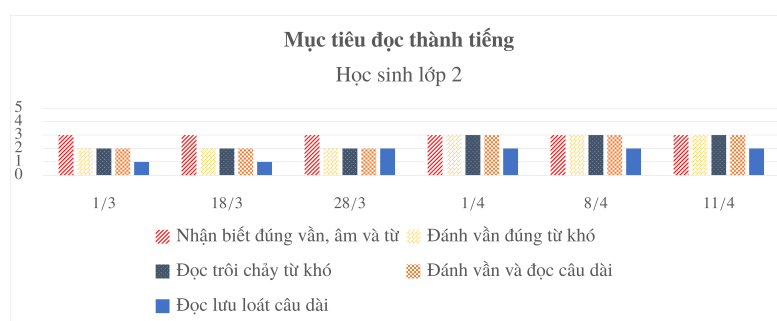
Luận án đề xuất và phát triển ứng dụng di động nhằm hỗ trợ trẻ em mắc chứng khó đọc. Kiến trúc của ứng dụng này bao gồm bốn khối nhiệm vụ chính hoạt động nối tiếp nhau: xử lý tài liệu đầu vào, khám phá đặc trưng của trẻ, hỗ trợ tương tác đa phương thức và kết xuất tài liệu đầu ra. Sơ đồ luồng dữ liệu cùng mối quan hệ chức năng giữa các khối được thể hiện trong Hình 4.10.



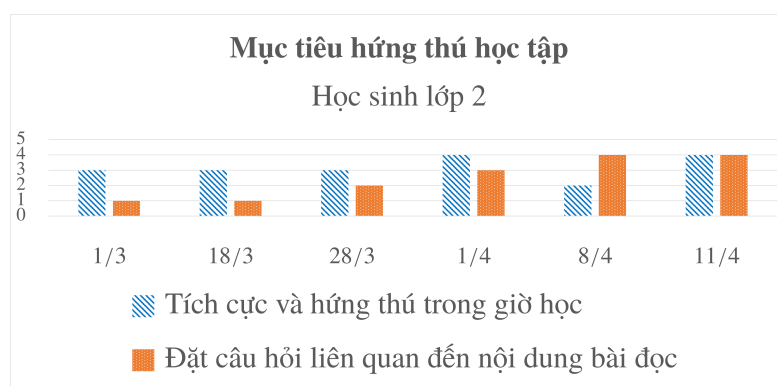
Hình 4.10: Mô hình kiến trúc tổng quan luồng dữ liệu thích ứng của hệ thống can thiệp ngôn ngữ.



Hình 4.11: Kết quả đạt được mục tiêu Đọc thành tiếng qua các buổi thực nghiệm – học sinh lớp 1

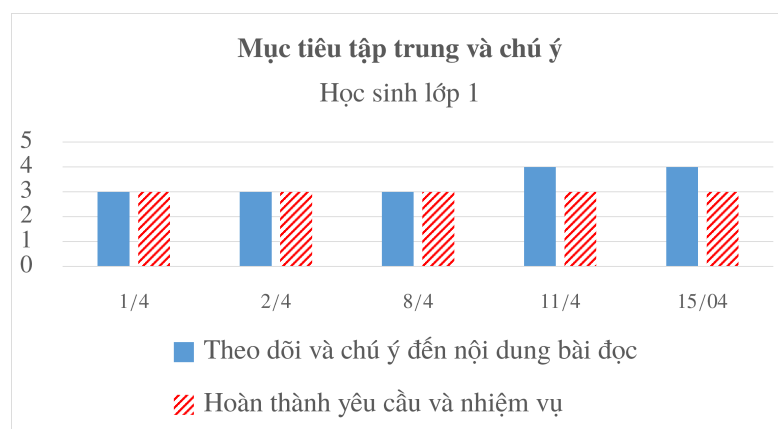


Hình 4.12: Kết quả đạt được mục tiêu Đọc thành tiếng qua các buổi thực nghiệm – học sinh lớp 2

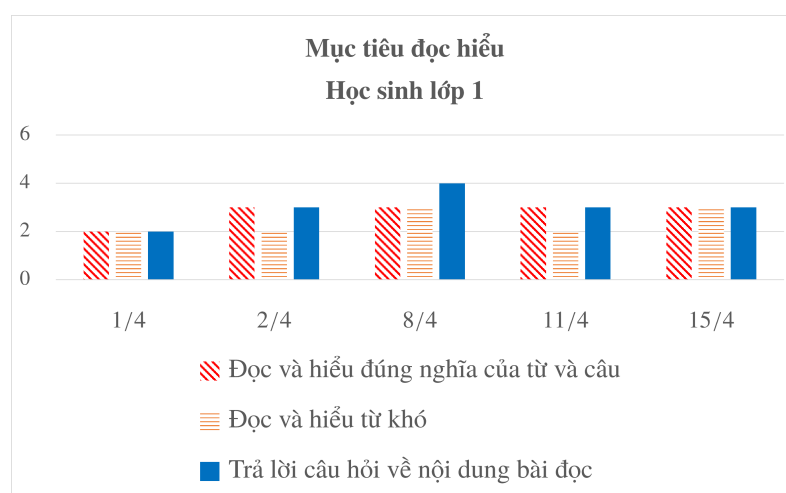


Hình 4.13: Kết quả đạt được mục tiêu hứng thú qua các buổi thực nghiệm – học sinh lớp 2

Nền tảng được thiết lập với hai luồng tương tác riêng biệt dành cho hai nhóm đối tượng:



Hình 4.14: Kết quả đạt được mục tiêu tập trung và chú ý qua các buổi thực nghiệm – học sinh lớp 1

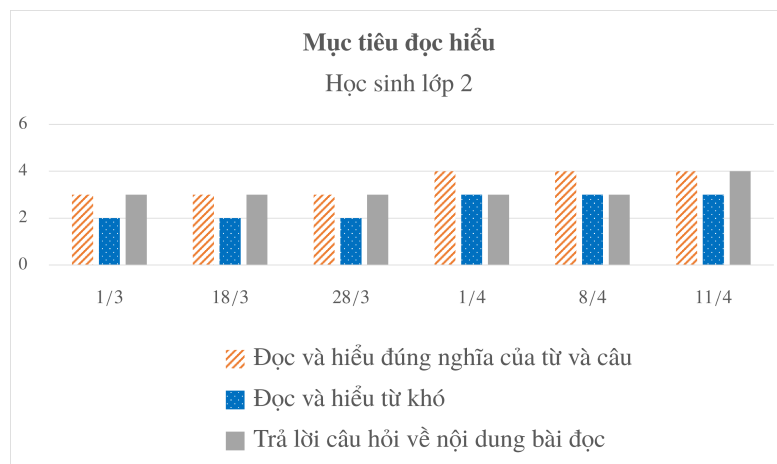


Hình 4.15: Kết quả đạt được mục tiêu đọc hiểu qua các buổi thực nghiệm – học sinh lớp 1

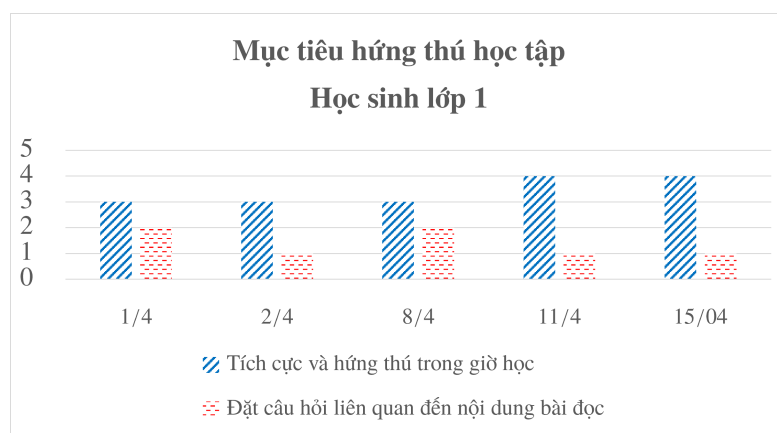
- Luồng tương tác của trẻ: trẻ trực tiếp thực hiện các nhiệm vụ rèn luyện kỹ năng đọc hiểu và tiếp nhận các cơ chế hỗ trợ thích ứng từ ứng dụng.
- Luồng tương tác của giáo viên: Nền tảng cung cấp chức năng để nhà chuyên môn thiết lập cấu hình bài học, khởi tạo học liệu số và theo dõi tiến trình thực hành của trẻ nhằm điều chỉnh can thiệp cho từng cá nhân.

Trong quá trình trẻ sử dụng ứng dụng liên tục ghi nhận các tương tác thông qua hành vi thị giác và kết quả trả lời câu hỏi. Chu trình phản hồi này được lưu vào thông tin cá nhân, đóng vai trò là căn cứ khách quan để giáo viên điều chỉnh cấu trúc bài học trong các phiên thực hành tiếp theo.

Đánh giá thực nghiệm trên nền tảng EzRead



Hình 4.16: Kết quả đạt được mục tiêu đọc hiểu qua các buổi thực nghiệm – học sinh lớp 2

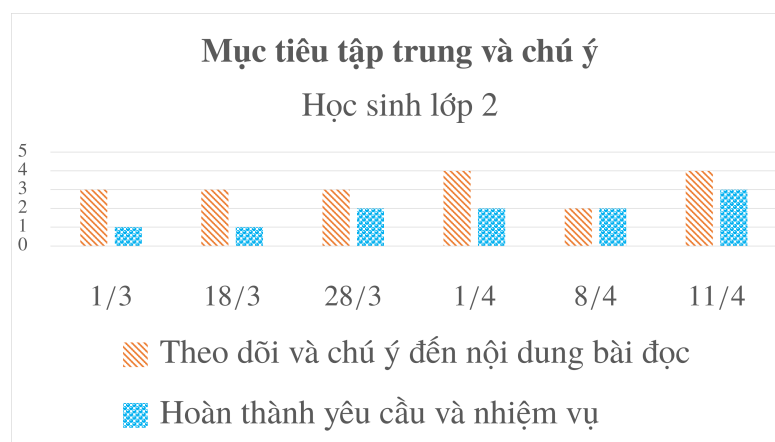


Hình 4.17: Kết quả đạt được mục tiêu hứng thú qua các buổi thực nghiệm – học sinh lớp 1

Quá trình thực nghiệm được triển khai theo mô hình can thiệp cá nhân hóa đối với hai học sinh tiểu học mắc chứng khó đọc. Các phiên can thiệp diễn ra với tần suất hai buổi mỗi tuần (thời lượng 30 đến 60 phút mỗi phiên) trên thiết bị máy tính bảng, dưới sự giám sát trực tiếp của giáo viên chuyên môn. Nguồn học liệu được số hóa và tích hợp vào nền tảng thông qua quy trình nhận dạng ký tự quang học (Optical Character Recognition - OCR).

Kết quả quan sát và dữ liệu thu thập cho thấy hệ thống các tính năng tùy chỉnh đem lại những chuyển biến tích cực đối với năng lực xử lý thông tin của trẻ, thể hiện qua ba khía cạnh chính:

- Tối ưu hóa không gian thị giác và năng lực đọc thành tiếng: Trẻ có xu hướng chủ động mở rộng cỡ chữ và khoảng cách ký tự ngay từ giai đoạn đầu tiếp cận hệ thống. Sự kết hợp giữa tính năng làm nổi bật từ



Hình 4.18: Kết quả đạt được mục tiêu tập trung và chú ý qua các buổi thực nghiệm – học sinh lớp 2

Bảng 4.4: Các thành phần tùy chỉnh hệ thống và mục tiêu hỗ trợ nhận thức

Thành phần tùy chỉnh trên nền tảng	Mục tiêu hỗ trợ nhận thức và hành vi
Cung cấp hiệu ứng làm nổi bật từ vựng theo tiến độ đọc	Định hướng ánh nhìn của trẻ bám sát đúng vị trí và nhận diện khoảng cách giữa các từ.
Chia nhỏ nội dung văn bản (<i>Luyện đọc theo phân đoạn và luyện đọc theo câu đơn</i>)	Giảm bớt áp lực tiếp nhận thông tin và khắc phục tình trạng rối mắt do quan sát mật độ chữ quá dày.
Kết hợp phản hồi bằng âm thanh và hình ảnh để xác nhận thao tác đúng	Duy trì sự chú ý và ổn định nhịp độ tương tác ổn định trong suốt quá trình thực hành.
Hiển thị đoạn văn gợi ý kèm theo hệ thống câu hỏi kiểm tra	Hỗ trợ trẻ đối chiếu dữ kiện thuận lợi và tìm kiếm câu trả lời nhanh chóng hơn.

vựng và hỗ trợ âm thanh hỗ trợ trẻ giảm thiểu tình trạng bỏ sót dòng, nhảy chữ, đồng thời tự nhận thức và điều chỉnh các lỗi phát âm sai (Hình 4.11, 4.12).

- Hỗ trợ năng lực đọc hiểu: Việc thiết lập các gợi ý trực quan song hành cùng hệ thống câu hỏi kiểm tra đóng vai trò *cơ chế giàn giáo nhận thức*⁹, hỗ trợ trẻ đối chiếu dữ kiện và phân tích nội dung bài đọc thuận lợi hơn (Hình 4.15, 4.16).

9. *Cơ chế giàn giáo nhận thức (Cognitive Scaffolding)*: là khái niệm do Wood, Bruner và Ross (1976) định hình dựa trên lý thuyết về Vùng phát triển gần (ZPD - Zone of Proximal Development) của Vygotsky. Thuật ngữ này chỉ các biện pháp hỗ trợ tạm thời từ hệ thống (như các gợi ý trực quan, chia nhỏ nhiệm vụ) nhằm giảm tải nhận thức và định hướng chú ý, giúp

- Gia tăng mức độ hứng thú và duy trì sự tập trung: Cơ chế phản hồi tích cực thông qua âm thanh và hình ảnh động viên thiết lập một môi trường tương tác thân thiện, hỗ trợ trẻ duy trì nhịp độ học tập ổn định và hạn chế các yếu tố gây xao nhãng (Hình 4.17, 4.18).

Dựa trên chuỗi phản hồi thực tế trong suốt tiến trình can thiệp, nền tảng EzRead liên tục được tinh chỉnh để đáp ứng nhu cầu của đối tượng sử dụng. Mối tương quan giữa các thành phần tùy chỉnh trên hệ thống và mục tiêu hỗ trợ nhận thức tương ứng được thống kê tại Bảng 4.4.

4.5 Mô hình suy diễn trạng thái nhận thức từ dữ liệu đa phương thức

Các nền tảng EzMath và EzRead thực hiện điều chỉnh lộ trình can thiệp dựa trên kết quả thao tác đúng hoặc sai của trẻ. Tuy nhiên, phương thức đánh giá nhị phân này hầu như chưa phản ánh đầy đủ tiến trình tư duy nội tại và các rào cản nhận thức của trẻ. Nhằm gia tăng mức độ thích ứng cho hệ thống, luận án đề xuất một mô hình ước lượng trạng thái nhận thức dựa trên luồng dữ liệu đa phương thức. Mô hình này đóng vai trò là phân hệ phân tích trung tâm, thực hiện nhiệm vụ ước lượng các trạng thái nhận thức thông qua dữ liệu tương tác giao diện và thời gian phản hồi, từ đó cung cấp cơ sở định lượng để xác định các lỗ hổng tri thức của trẻ.

4.5.1 Kiến trúc vòng lặp kép dựa trên đồ thị phụ thuộc phân tích nhiệm vụ

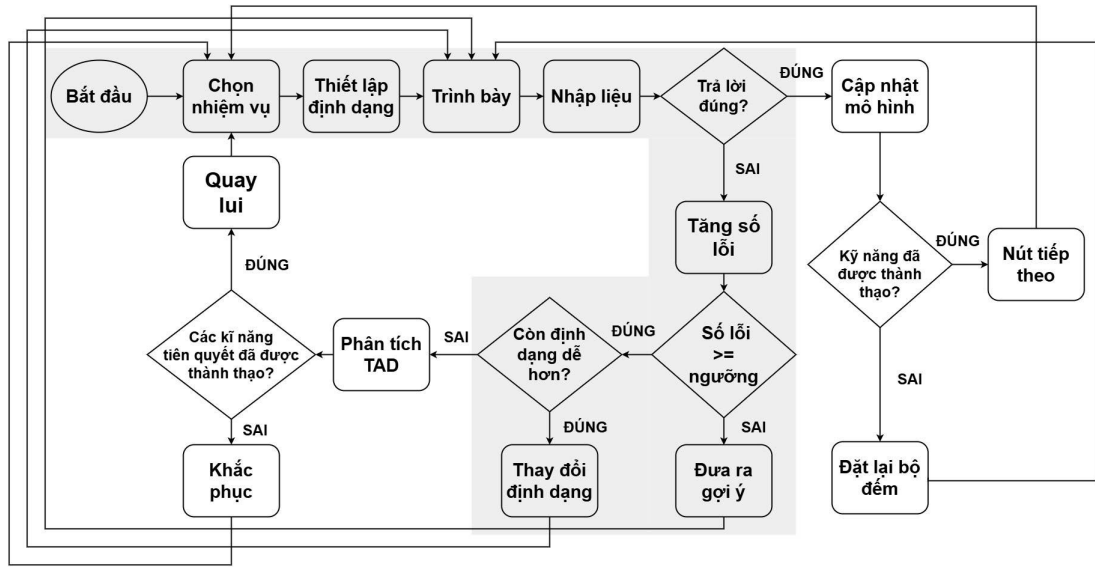
Các hệ thống hỗ trợ học tập truyền thống thường xử lý lỗi theo mô hình tuyến tính: khi trẻ chưa hoàn thành yêu cầu tại một kỹ năng, hệ thống sẽ lặp lại nội dung tại chính kỹ năng đó hoặc chuyển hướng sang kỹ năng lân cận. Cách tiếp cận này thể hiện hạn chế trong việc định vị vùng kiến thức thiếu hụt, vì vị trí xuất hiện lỗi không nhất thiết đồng nhất với nguyên nhân chính của sự khiếm khuyết nhận thức [168].

Để khắc phục hạn chế này, nghiên cứu [168] áp dụng cấu trúc TAD, ký hiệu $G_{TAD} = (V, E)$, làm nền tảng biểu diễn tri thức. Trong đó, V là tập

người học hoàn thành các nhiệm vụ mà họ chưa đủ năng lực để thực hiện độc lập. Khi năng lực của trẻ tăng lên, các hỗ trợ này sẽ được rút dần [166, 167].

hợp các nút kỹ năng được phân bổ theo bốn tầng phụ thuộc và E là tập các cạnh có hướng biểu diễn mối quan hệ tiên quyết giữa các kỹ năng. Toàn bộ tiến trình suy diễn trên cấu trúc này được đặt trong khung mạng Bayes động: tại mỗi bước thời gian t , mô hình cập nhật xác suất hậu nghiệm $P(K_t = 1 \mid O_t)$ về trạng thái làm chủ kiến thức ẩn của từng nút kỹ năng, dựa trên chuỗi bằng chứng quan sát gồm tính chính xác và thời gian phản hồi. Các minh chứng thực nghiệm từ nghiên cứu [169] cho thấy cách tiếp cận DBN có khả năng phân tách hai dạng hồ sơ nhận thức (lỗi phân biệt âm vị và lỗi cô lập âm vị cuối), từ đó khởi tạo các lộ trình can thiệp mang tính cá nhân hóa, nguyên lý này được kế thừa và mở rộng trong nghiên cứu [168] để định vị nguyên nhân theo cấu trúc thứ bậc kỹ năng của đồ thị G_{TAD} . Dựa trên nền tảng G_{TAD} , một kiến trúc vòng lặp kép được thiết kế nhằm xử lý hiệu quả với dữ liệu tương tác hiện tại, đồng thời sẵn sàng mở rộng để tích hợp dữ liệu sinh trắc học thị giác. Kiến trúc này bao gồm:

Vòng lặp trong: Phân hệ này chịu trách nhiệm hiệu chỉnh các thông số hiển thị và học liệu dựa trên luồng dữ liệu tương tác tức thời của trẻ. Cơ chế vận hành được xây dựng trên cấu trúc cây quyết định, phân loại phản hồi theo hai chiều dữ liệu: tính chính xác của kết quả và thời gian thực hiện. Khi ghi nhận lỗi thao tác, hệ thống kích hoạt cơ chế thăm dò ngữ cảnh [170] thông qua việc cung cấp các nhiệm vụ thuộc chủ đề khác nhưng bảo toàn cấu trúc mục tiêu nhiệm vụ. Quá trình này giúp hệ thống phân tách lỗi do tình huống xao nhãng khỏi sự thiếu hụt kiến thức nền tảng. Khi một điểm yếu được xác nhận qua nhiều ngữ cảnh kiểm chứng, hệ thống tiến hành thay đổi định dạng nhiệm vụ [170] (Hình 4.19). Tùy thuộc vào đặc điểm lỗi, phương thức tương tác có thể được chuyển đổi từ nhập liệu văn bản sang thao tác kéo thả, hoặc điều chỉnh các thuộc tính hiển thị nhằm giảm tải hạn chế tri giác. Mục tiêu trọng tâm của vòng lặp trong là giảm thiểu các yếu tố nhiễu do rào cản giao diện, ngữ cảnh qua đó cung cấp nguồn dữ liệu đầu vào duy trì độ tin cậy cho vòng lặp ngoài.

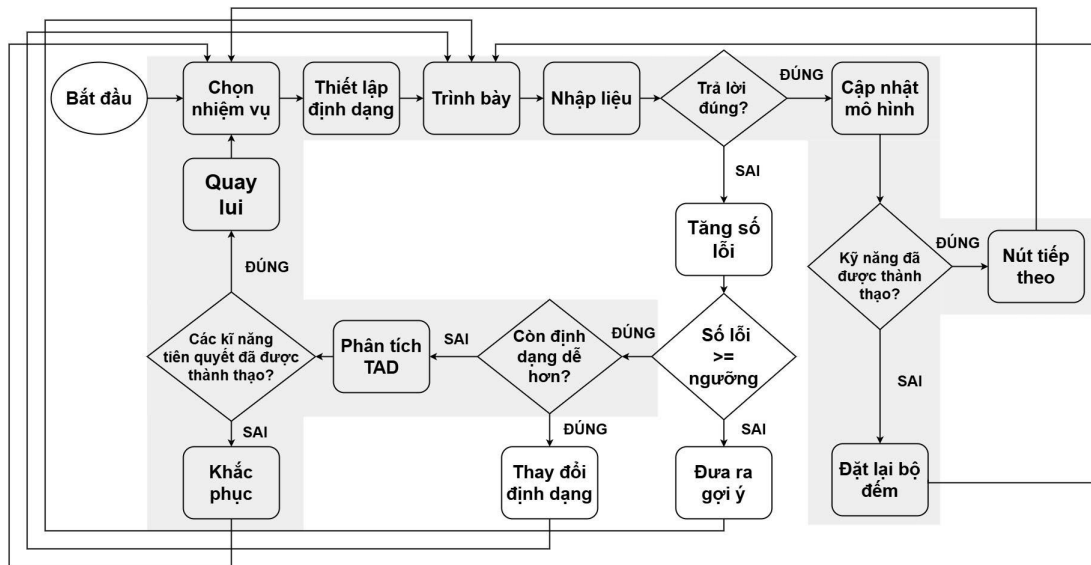


Hình 4.19: Vòng lặp trong: Cơ chế hạ định dạng

Vòng lặp ngoài: Phân hệ này điều hướng lộ trình can thiệp dài hạn dựa trên cấu trúc đồ thị G_{TAD} . Khi vòng lặp trong xác nhận nguyên nhân lỗi xuất phát từ hệ thống kiến thức, vòng lặp ngoài sẽ khởi chạy thuật toán truy vết ngược trên đồ thị thông qua cơ chế duyệt các cạnh hướng lên [168] (Hình 4.20). Do mỗi bước truy vết đánh giá trạng thái làm chủ của một nút kỹ năng đơn lẻ dựa trên chuỗi phản hồi tại chính nút đó, tiến trình cập nhật $P(K_t = 1 | O_t)$ được hiện thực hóa dưới dạng mô hình truy vết kiến thức Bayes (BKT) với bốn tham số đặc trưng: $P(L_0)$ (xác suất làm chủ ban đầu), $P(L)$ (tỉ lệ học), $P(G)$ (tỉ lệ đoán) và $P(S)$ (tỉ lệ trượt) [171]. Tại mỗi bước tính toán, thuật toán truy xuất tập nút cha $\{p\}$ của nút hiện tại $S_{current}$ và đánh giá xác suất làm chủ tri thức $P(K_t = 1)$ của từng nút cha theo cơ chế này. Hệ thống ưu tiên mở rộng nhánh đối với các nút cha có xác suất $P(K_t = 1)$ thấp hoặc có chu kỳ bỏ trống luyện tập dài hơn. Khi phát hiện một nút cha chưa đạt ngưỡng thành thạo, thuật toán thực hiện gọi đệ quy để truy vết cho đến khi định vị được nút tiên quyết nền tảng bị thiếu hụt. Lộ trình can thiệp sau đó được tái thiết lập bắt đầu từ nút tiên quyết này, đảm bảo chiến lược hỗ trợ giải quyết đúng vấn đề thay vì chỉ xử lý các biểu hiện lỗi bề mặt.

Đồng thời, nguyên lý phân tách lỗi hiệu năng khỏi lỗi năng lực bằng cách thay đổi định dạng hoặc ngữ cảnh câu hỏi, đã được kiểm chứng thực nghiệm trong nghiên cứu [170], được tích hợp trực tiếp vào cơ chế thăm dò ngữ cảnh của vòng lặp trong [168]. Tương tự, cơ chế chẩn đoán dựa trên

xác suất hậu nghiệm nhằm phân tách các hồ sơ nhận thức khác nhau, đã được kiểm chứng thực nghiệm trong nghiên cứu [169], được kế thừa và mở rộng theo cấu trúc thứ bậc kỹ năng của đồ thị G_{TAD} trong vòng lặp ngoài [168]. Như vậy, kiến trúc vòng lặp kép [168] đóng vai trò tổng hợp hai cơ chế chẩn đoán này: vòng lặp trong kế thừa nguyên lý thăm dò ngữ cảnh từ [170], còn vòng lặp ngoài kế thừa nguyên lý suy diễn xác suất theo thứ bậc kỹ năng từ [169].



Hình 4.20: Vòng lặp ngoài: Truy vết ngược trên TAD

Thiết kế thực nghiệm trường hợp đơn: Khả năng giảm thiểu rào cản thao tác và định vị lỗi hồng nền tảng

Nhằm kiểm tra khả năng hoạt động độc lập của từng phân hệ trong kiến trúc vòng lặp kép, phần này trình bày hai quá trình phân tích tình huống riêng biệt: một phân tích nhằm đánh giá cơ chế thăm dò ngữ cảnh của vòng lặp trong (kế thừa từ nghiên cứu [170]), và một phân tích nhằm kiểm chứng khả năng phân loại đặc trưng của vòng lặp ngoài (kế thừa từ nghiên cứu [169]). Hai quá trình này được thực hiện bằng cách tái phân tích dữ liệu từ các nhóm đối tượng và bối cảnh khác nhau của từng nghiên cứu gốc, qua đó làm rõ giá trị đóng góp riêng của từng phân hệ trước khi kết hợp chúng vào một kiến trúc thống nhất.

Thực nghiệm vòng lặp trong: Phân tách lỗi hiệu năng khỏi lỗi năng lực

Thực nghiệm này được thực hiện trong khuôn khổ nghiên cứu [170], với sự tham gia của 5 học sinh lớp một (tuổi trung bình 6) sử dụng tiếng Việt

làm ngôn ngữ chính, với các mức độ rào cản học tập khác biệt. Quy trình thực nghiệm kéo dài 4 tuần (2 đến 3 buổi/tuần) trong môi trường lớp học có kiểm soát. Mỗi phiên tương tác kéo dài 20 đến 30 phút, bao gồm ba giai đoạn: phân tích lỗi qua bài đọc được ghi âm, thiết lập ưu tiên mục tiêu can thiệp và luyện tập có mục tiêu với bài tập được thuật toán lựa chọn trong vùng phát triển gần của trẻ.

Trong một bài tập đọc, hệ thống ghi nhận một chuỗi tương tác đặc trưng với học sinh HS03, thể hiện cơ chế đánh giá vi mô của thuật toán thích ứng. Khi được yêu cầu đọc câu "Ánh nắng chiếu khắp mọi nơi", hệ thống ghi nhận hai lỗi phát âm liên quan đến các âm gần giống: từ mục tiêu "nắng" bị đọc sai thành "năng" (lỗi âm /ă/) và từ mục tiêu "khắp" bị đọc sai thành "khắc" (lỗi vần /ăp/).

Thay vì lập tức ghi nhận sự thiếu hụt kỹ năng và tự động giảm độ khó của bài tập, hệ thống kích hoạt cơ chế thăm dò ngữ cảnh. Một câu văn thuộc chủ đề khác là *Bạn Hằng sắp đọc bài trước lớp* được cung cấp. Câu văn này được lựa chọn có chủ đích vì chứa hai từ *Hằng* và *sắp*, nhằm trực tiếp kiểm tra lại hai cấu trúc âm vị bị lỗi trước đó trong một ngữ cảnh mới. Dữ liệu cho thấy đối tượng HS03 đọc chính xác toàn bộ câu thứ hai. Dựa trên cơ sở này, hệ thống xác định các lỗi ban đầu là lỗi hiệu năng tạm thời chứ không phản ánh sự thiếu hụt về năng lực cốt lõi. Nhờ đó, tham số năng lực của trẻ được giữ nguyên và lộ trình can thiệp tiếp tục mà không kích hoạt vòng lặp củng cố không cần thiết.

Bảng 4.5: Kết quả đánh giá trước và sau can thiệp của 5 học sinh (thực nghiệm vòng lặp trong)

Mã học sinh	Trước can thiệp (/100)	Sau can thiệp (/100)	Mức cải thiện
HS1	49	87	38
HS2	45	77	32
HS3	28	77	49
HS4	62	80	18
HS5	71	79	8

Đối chiếu với toàn bộ mẫu thực nghiệm, dữ liệu tổng hợp từ bài đánh giá trước và sau can thiệp của 5 trẻ (Bảng 4.5) cho thấy mức cải thiện độ chính xác nhất quán ở tất cả các trường hợp, dao động từ 8 đến 49 điểm trên thang 100.

Kết quả của HS03, đặt trong bối cảnh mức cải thiện chung của toàn mẫu, cho thấy cơ chế thăm dò ngữ cảnh không chỉ ngăn chặn việc hạ thấp độ khó không cần thiết ở cấp độ từng tương tác, mà còn góp phần duy trì tính chính xác của lộ trình can thiệp tổng thể. Đây là nguyên lý được vòng lặp trong của kiến trúc đề xuất kế thừa trực tiếp [168].

Thực nghiệm vòng lặp ngoài: Khả năng chẩn đoán phân biệt

Đối với thực nghiệm vòng lặp ngoài, kết quả chẩn đoán phân biệt trên hai trẻ có khó khăn giải mã từ (Bảng 4.6, Bảng 4.7, trình bày ở mục 4.5.2.2) cho thấy mô hình có khả năng phân tách hai hồ sơ nhận thức khác biệt, gồm thiếu hụt âm cuối và thiếu hụt phân biệt âm vị, từ cùng một bộ dữ liệu phản hồi đầu vào, qua đó khởi tạo hai lộ trình can thiệp cá nhân hóa tương ứng. Đây là nguyên lý được vòng lặp ngoài của kiến trúc đề xuất kế thừa và mở rộng theo cấu trúc thứ bậc kỹ năng của đồ thị G_{TAD} [168].

4.5.2 Mạng Bayes động và mô hình hóa nhận thức

Nhằm thiết lập một mô hình đối chứng quá trình thực nghiệm, phần này trình bày cấu trúc toán học của mạng Bayes động truyền thống, độc lập với các tham số thị giác. Trong lĩnh vực khoa học máy tính và công nghệ giáo dục, kiến trúc này được biết đến thông qua mô hình Truy vết kiến thức Bayes (Bayesian Knowledge Tracing - BKT) do Corbett và Anderson (1995) [153] đề xuất. Đặc điểm chính của mô hình BKT là cơ chế suy diễn dựa trên luồng dữ liệu tương tác nhị phân; cụ thể, hệ thống ghi nhận và xử lý kết quả thao tác của trẻ trên giao diện theo hai trạng thái đối lập: thực hiện đúng hoặc thực hiện sai nhiệm vụ.

4.5.2.1 Các biến trạng thái và tham số xác suất

Mô hình toán học của một mạng DBN truyền thống đối với một kỹ năng cụ thể tại bước thời gian t được xác định dựa trên hai tầng biến ngẫu nhiên không phụ thuộc vào dữ liệu sinh trắc học [153, 171]:

- Trạng thái ẩn (K_t): Biểu diễn mức độ làm chủ kiến thức ẩn của trẻ tại thời điểm t . Miền giá trị của trạng thái mang tính chất nhị phân: $K_t \in \{0, 1\}$, với $K_t = 1$ đại diện cho trạng thái đã làm chủ kỹ năng, và $K_t = 0$ là chưa làm chủ.

- Biến quan sát (O_t): Biểu diễn hành vi đầu ra thực tế của trẻ khi thực hiện nhiệm vụ tại thời điểm t . Trong cấu trúc mô hình, biến này là kết quả làm bài: $O_t \in \{0, 1\}$ (0: Trả lời sai, 1: Trả lời đúng) [171].

Tiến trình cập nhật và suy diễn xác suất của mạng DBN được quyết định bởi một bộ 4 tham số xác suất tĩnh cố định [48]:

- $P(K_0)$ (Xác suất tiên nghiệm): Xác suất trẻ đã làm chủ được kĩ năng trước khi tương tác với hệ thống.
- $P(T)$ (Tỷ lệ học): Xác suất trẻ chuyển từ trạng thái chưa làm chủ kiến thức ($K_{t-1} = 0$) sang đã làm chủ kiến thức ($K_t = 1$) sau một bước tương tác nhiệm vụ.
- $P(G)$ (Tỷ lệ đoán mò - Guessing): Xác suất trẻ chưa làm chủ kiến thức ($K_t = 0$) nhưng vẫn đưa ra câu trả lời đúng ($O_t = 1$).
- $P(S)$ (Tỷ lệ trượt - Slipping): Xác suất trẻ đã làm chủ kiến thức ($K_t = 1$) nhưng vẫn đưa ra câu trả lời sai ($O_t = 0$) do bất cẩn hoặc rào cản tương tác vật lý.

4.5.2.2 Thuật toán cập nhật suy diễn Bayes

Tại mỗi bước thời gian t , ngay sau khi hệ thống ghi nhận kết quả hành vi O_t , xác suất hậu nghiệm về trạng thái làm chủ kiến thức ẩn $P(K_t = 1 | O_t)$ được tính toán cục bộ thông qua việc áp dụng trực tiếp định lý Bayes dựa trên hai kịch bản đầu ra [48]:

Trường hợp 1: trẻ đưa ra câu trả lời đúng ($O_t = 1$):

$$P(K_t = 1 | O_t = 1) = \frac{P(K_{t-1} = 1) \cdot (1 - P(S))}{P(K_{t-1} = 1) \cdot (1 - P(S)) + (1 - P(K_{t-1} = 1)) \cdot P(G)} \quad (4.6)$$

Trường hợp 2: trẻ đưa ra câu trả lời sai ($O_t = 0$):

$$P(K_t = 1 | O_t = 0) = \frac{P(K_{t-1} = 1) \cdot P(S)}{P(K_{t-1} = 1) \cdot P(S) + (1 - P(K_{t-1} = 1)) \cdot (1 - P(G))} \quad (4.7)$$

Sau khâu cập nhật hậu nghiệm, mạng DBN thực hiện bước dự báo chuyển trạng thái để ước lượng xác suất tiên nghiệm cho chu trình tương tác tiếp theo tại thời điểm $t + 1$:

$$P(K_{t+1} = 1) = P(K_t = 1 | O_t) + [1 - P(K_t = 1 | O_t)] \cdot P(T) \quad (4.8)$$

Trong đó, tỷ lệ quên của trẻ theo thời gian ngắn được giả định bằng hằng số không ($P(F) = 0$) để bảo đảm tính hội tụ tuyến tính cho thuật toán.

Thực nghiệm trường hợp đơn: Khả năng chẩn đoán phân biệt

Nhằm đánh giá năng lực chẩn đoán phân biệt của mô hình DBN, một thực nghiệm trường hợp đơn được tiến hành trên hai trẻ ($N = 2$) có khó khăn trong kỹ năng giải mã từ. Quy trình thực nghiệm gồm ba bước: đánh giá trước can thiệp, can thiệp thích ứng qua hệ thống ITS với nội dung cá nhân hóa theo mô hình chẩn đoán thời gian thực và đánh giá sau can thiệp để đo mức cải thiện.

Tại bước đánh giá đầu vào, hai trẻ đọc thành tiếng ba từ mục tiêu "trâu", "học", "ngã" (Bảng 4.6), cho thấy dạng lỗi và thời gian phản hồi khác nhau giữa hai trẻ.

Bảng 4.6: Phản hồi trong bài đánh giá trước can thiệp

Từ	Trẻ 1	Thời gian (giây)	Trẻ 2	Thời gian (giây)
trâu	tâu	5,1	trâu	8,2
học	họ	3,1	hóc	3,5
ngã	ngả	5,5	ga	4,8

Dựa trên dữ liệu này, mô hình DBN sinh ra hai hồ sơ chẩn đoán khác biệt và khởi tạo hai lộ trình can thiệp riêng.

Trường hợp 1 (Trẻ 1), thiếu hụt âm cuối: Phản hồi sai nhưng nhanh (3,1 giây) đối với từ "học" được DBN xác định là biểu hiện của thiếu hụt kiến thức thực sự, không phải lỗi ngẫu nhiên ($P(S)$ thấp), quy về sự chưa làm chủ kỹ năng nhận diện âm cuối. Lộ trình can thiệp ba giai đoạn được khởi tạo: (1) liên kết thính giác – thị giác giữa âm vị /c/ và chữ viết tương ứng; (2) nhận thức vị trí phụ âm cuối qua sắp xếp khối âm vị; (3) tự động hóa qua lặp lại và đọc độc lập từ mục tiêu.

Trường hợp 2 (Trẻ 2), thiếu hụt phân biệt âm vị: Hồ sơ chẩn đoán cho thấy thiếu hụt liên quan đến phân biệt thính giác giữa các âm vị gần giống (ví dụ /ng/ và /g/) và thanh điệu. Lộ trình can thiệp gồm: (1) phân biệt cặp âm tối thiểu (ví dụ "ga" và "nga"); (2) nhận diện từ bắt đầu bằng âm mục tiêu; (3) tự động hóa qua đọc thành tiếng các cặp từ dễ nhầm lẫn. Mặc dù độ chính xác cải thiện ở giai đoạn 1, tốc độ phản hồi vẫn chậm, cho thấy

$P(L)$ chưa đạt ngưỡng làm chủ; hệ thống tự động bổ sung bài luyện tập củng cố trước khi chuyển giai đoạn.

Giá trị $P(L)$ (Bảng 4.7) khác nhau giữa hai hồ sơ nhận thức, phù hợp với hai chẩn đoán đã nêu. Tại bước đánh giá sau can thiệp, hai trẻ đọc chính xác các từ mục tiêu với tốc độ phản hồi được cải thiện, xác nhận phương pháp dựa trên mô hình DBN có khả năng thích ứng linh hoạt và tùy biến nội dung can thiệp theo đặc điểm nhận thức của từng trẻ.

Bảng 4.7: Giá trị $P(L)$ của các kỹ năng tiên quyết

Kỹ năng	Trẻ 1	Trẻ 2
1.14 Phân biệt cặp âm vị	0,2	0,1
1.19 Phân biệt thanh điệu	0,3	0,2
1.26 Nhận diện cấu trúc ngôn ngữ	0,1	0,9

4.5.2.3 Hạn chế về hiện tượng bất định danh tham số

Mặc dù cấu trúc toán học của mạng DBN truyền thống được xác lập rõ ràng, mô hình này vẫn tồn tại hạn chế trong quá trình ước lượng tham số, đặc trưng bởi hiện tượng *bất định danh*¹⁰. Do hàm mục tiêu xác suất (likelihood) chỉ được xác định dựa trên luồng dữ liệu tương tác nhị phân, không gian nghiệm tồn tại nhiều bộ tham số ($P(G), P(S)$) khác nhau cùng hội tụ về một giá trị (ví dụ: hệ thống không thể phân định chính xác giữa trường hợp trẻ có năng lực cao nhưng thao tác lỗi và trường hợp trẻ thực sự có mức năng lực thấp) [172].

Đặc điểm toán học này đặt ra yêu cầu tất yếu về việc mở rộng bộ dữ liệu đầu vào thay vì chỉ phụ thuộc vào dữ liệu nhị phân. Do đó, luận án đề xuất một cấu trúc mạng DBN cải tiến, trong đó các tham số xác suất quan sát $P(S)$ và $P(G)$ được tái cấu trúc dưới dạng các hàm số biến thiên động, phụ thuộc trực tiếp vào luồng dữ liệu sinh trắc học thị giác.

10. *Bất định danh tham số (Parameter Non-identifiability)*: Trong thống kê, hiện tượng này xảy ra khi nhiều nhóm nguyên nhân khác nhau lại cùng tạo ra một kết quả tính toán giống hệt nhau. Khi áp dụng vào các mô hình ước lượng năng lực, điều này có nghĩa là hệ thống không thể xác định chính xác năng lực thực sự của trẻ nếu chỉ quan sát một lượng dữ liệu bề mặt quá hẹp (ví dụ: chỉ nhìn vào kết quả đúng/sai). Sự hạn chế này khiến hệ thống không thể phân biệt được trẻ làm sai do thiếu kiến thức, hay làm đúng nhờ đoán mò, một vấn đề nền tảng đã được Molenaar (1985) và Beck & Chang (2007) phân tích chi tiết [172, 173].

4.5.2.4 Mạng Bayes động cải tiến tích hợp hàm xác suất quan sát động

Nhằm giải quyết rào cản về hiện tượng bất định danh của mô hình BKT truyền thống, luận án đề xuất cấu trúc mạng Bayes động cải tiến. Mô hình này được định nghĩa một cách chặt chẽ thông qua bộ năm thành phần toán học [153, 174]:

$$\mathcal{M}_{DBN} = \langle \mathcal{K}, \mathcal{O}, \mathcal{X}, \mathcal{T}, \mathcal{E} \rangle \quad (4.9)$$

Trong đó:

- $\mathcal{K}$ là không gian trạng thái ẩn biểu diễn mức độ làm chủ kiến thức, $\mathcal{K} \in \{0, 1\}$.
- $\mathcal{O}$ là không gian biến quan sát nhị phân thể hiện kết quả tương tác đúng hoặc sai, $\mathcal{O} \in \{0, 1\}$.
- $\mathcal{X}$ là vectơ đặc trưng hành vi thị giác đầu vào tại mỗi bước thời gian.
- $\mathcal{T}$ là ma trận xác suất chuyển trạng thái kiến thức giữa các thời điểm liên tiếp.
- $\mathcal{E}$ là tập hợp các hàm xác suất quan sát xác định mối quan hệ giữa trạng thái ẩn và biến quan sát.

Trong cấu trúc đề xuất, các tham số xác suất quan sát không duy trì ở dạng hằng số tĩnh mà được chuyển đổi thành cấu trúc hàm xác suất quan sát động phụ thuộc vào vectơ đặc trưng thị giác đầu vào $\mathcal{X}_t = [F_t, C_t]^T$. Trong đó, F_t đại diện cho thời lượng điểm nhìn cố định và C_t đại diện cho số lượng điểm nhìn cố định, được đưa vào phân hệ tính toán dưới dạng các biến số định lượng thu thập từ quá trình tương tác. Xác suất trượt $P(S_t)$ và xác suất phỏng đoán $P(G_t)$ được tính toán liên tục thông qua hàm phân phối Logistic [175]:

$$P(S_t) = \frac{1}{1 + e^{-(\omega_1 F_t + \omega_2 C_t + \beta_S)}} \quad (4.10)$$

$$P(G_t) = \frac{1}{1 + e^{-(\omega_3 F_t + \omega_4 C_t + \beta_G)}} \quad (4.11)$$

Các trọng số ω và hệ số chặn β cần được ước lượng và tinh chỉnh thông qua các thuật toán học máy trên bộ dữ liệu sinh trắc học thực nghiệm.

Cấu trúc hàm số này hỗ trợ hệ thống điều chỉnh xác suất sai số dựa trên các biểu hiện hành vi được ghi nhận [171]. Khi giá trị thời lượng điểm nhìn cố định tăng cao, hàm phân phối sẽ điều chỉnh tăng giá trị xác suất trượt $P(S_t)$ tương ứng. Sự thay đổi này cung cấp cơ sở để thuật toán Forward suy diễn rằng kết quả trả lời sai có thể xuất phát từ yếu tố phân tán sự chú ý thị giác thay vì do thiếu hụt năng lực kiến thức, từ đó tiến hành cập nhật xác suất hậu nghiệm $P(K_t = 1 \mid O_t)$ [153].

4.5.2.5 Phân tích độ phức tạp và tính hội tụ

Tính đúng đắn và sự hội tụ của cơ chế truy vết ngược

Cơ chế truy vết ngược được thực thi trên đồ thị nghịch đảo $G^{-1} = (V, E^{-1})$. Do đồ thị gốc G là đồ thị có hướng không chu trình (Directed Acyclic Graph - DAG) [176], đồ thị nghịch đảo G^{-1} tiếp tục kế thừa đặc tính không chu trình và duy trì thứ tự tô-pô. Đặc tính này đảm bảo tính dừng của thuật toán duyệt đồ thị theo chiều sâu. Xét ở giới hạn cận trên, thuật toán sẽ hội tụ với số bước đệ quy không vượt quá N , với N là tổng số nút kỹ năng trong hệ thống.

Định hướng giải quyết hiện tượng bất định danh của mô hình DBN cải tiến

Về phương diện toán học, việc thiết lập các tham số xác suất quan sát dưới dạng hàm Logistic phụ thuộc vào vectơ đặc trưng thị giác $\mathcal{X}_t$ giúp duy trì ma trận hiệp phương sai ở trạng thái không suy biến. Dựa trên cơ sở đó, ma trận thông tin Fisher đạt giá trị xác định dương, hỗ trợ bộ tham số hội tụ về một cực đại toàn cục và xử lý hiệu quả hiện tượng bất định danh. Xét về mặt hiệu năng, độ phức tạp tính toán trong quá trình cập nhật xác suất Bayes tại từng bước thời gian được tối ưu ở mức $\mathcal{O}(1)$, duy trì tính khả thi để tích hợp thuật toán vào các nền tảng tương tác thời gian thực [177]. Dù vậy, trong giới hạn hiện tại, cấu trúc tính toán này đóng vai trò là khung lý thuyết định hướng, tạo tiền đề cho các nghiên cứu thực chứng sâu hơn khi hạ tầng thiết bị thu thập dữ liệu sinh trắc học trên nhóm đối tượng được triển khai đồng bộ.

4.6 Thảo luận chung

Mặc dù thiết kế SCED kết hợp thống kê Tau-U cho thấy hiệu quả đối với nhóm trẻ RLPTK, quy mô mẫu nhỏ còn là một hạn chế. Thiết kế SCED khắc phục điểm này bằng cách yêu cầu dữ liệu ổn định ở pha cơ sở (A) nhằm xác thực việc đo lường lặp lại không tạo ra sự tiến bộ tự nhiên, giúp giảm thiểu ảnh hưởng của hiệu ứng thực hành.

Trong tương lai, việc thu thập dữ liệu sinh trắc học của trẻ đòi hỏi các tiêu chuẩn về an toàn thông tin. Kiến trúc hệ thống được cấu hình tuân thủ hai nguyên tắc sau:

- Ẩn danh hóa dữ liệu tại biên: Dữ liệu hình ảnh và video từ cảm biến được xử lý cục bộ trên thiết bị để chuyển đổi thành ma trận tọa độ không gian - thời gian (x, y, t) , giúp loại bỏ thông tin nhận dạng cá nhân (PII) trước khi truyền về máy chủ.
- Phân quyền kiểm soát truy cập: Toàn bộ dữ liệu hành vi thị giác được mã hóa và áp dụng cơ chế phân quyền, giới hạn quyền tiếp cận cấu phần dữ liệu này chỉ cho chuyên gia can thiệp và phụ huynh.

Kết chương

Chương này đã trình bày về các phương pháp đề xuất nhằm hỗ trợ can thiệp cá nhân hóa cho nhóm trẻ rối loạn phát triển, giải quyết tính đa dạng trong các rào cản nhận thức đặc thù. Luận án đề xuất một kiến trúc hệ thống phân lớp theo định dạng mô-đun.

Tổng kết lại, chương này mang lại hai đóng góp chính: đề xuất các phương pháp hỗ trợ can thiệp thực tiễn và xây dựng quy trình mô hình hóa tri thức.

- Đề xuất các phương pháp hỗ trợ can thiệp: Luận án được triển khai theo hai hướng tiếp cận song song. Thứ nhất, đối với hoạt động học tập, các phần mềm EzMath và EzRead được thiết kế để tự động tinh chỉnh giao diện, giúp giảm bớt các yếu tố gây xao nhãng và tránh tình trạng trẻ bị quá tải thông tin. Thứ hai, luận án ứng dụng công nghệ theo dõi chuyển động mắt kết hợp với hệ cơ sở luật để hỗ trợ can thiệp trực tiếp cho nhóm trẻ rối loạn phổ tự kỉ. Phân tích thống kê cho thấy

phương pháp này đạt độ tin cậy rất cao (chỉ số Tau-U đạt 0,93), cho thấy hiệu quả trong việc giúp trẻ duy trì ổn định các kỹ năng sau can thiệp.

- Đề xuất quy trình mô hình hóa tri thức: Về mặt phương pháp luận, luận án sử dụng mạng Bayes động (DBN) và kiến trúc vòng lặp kép để liên tục theo dõi và ước lượng mức độ nắm vững kiến thức của trẻ. Thay vì chỉ đánh giá bề mặt qua việc trẻ làm đúng hay sai, quy trình mô hình hóa này giúp hệ thống hiểu rõ nền tảng nhận thức của trẻ. Đây chính là tiền đề thuật toán quan trọng để xây dựng các lộ trình giáo dục cá nhân hóa, phù hợp với đặc thù phát triển của từng trẻ.

Nội dung về các nền tảng thực thi và khung thuật toán trong chương này đã được công bố thông qua các công trình:

a. Các công trình đã được công bố/đăng tải

- **Thi-Quynh-Hoa Nguyen**, Thi-Cam-Huong Nguyen, Thanh-Ha Le, and Thi-Duyen Ngo. EzMath: Developing Digital Methods, a Website-Based Approach to Support Teaching Mathematics to First-Grade Vietnamese Children with Dyscalculia. In: *Journal of Science Natural Science*, 2025, 70(3), pp. 83–94. DOI: <https://doi.org/10.18173/2354-1059.2025-0039>.
- **Thi-Quynh-Hoa Nguyen**, Thi-Xuan-Thu Le, Thi-Cam-Huong Nguyen, Thanh-Ha Le, and Thi-Duyen Ngo. EzRead: A Mobile-Based Approach for Vietnamese Children with Dyslexia. In: *Proceedings of the 2025 7th Asia Pacific Information Technology Conference (APIT '25)*, Association for Computing Machinery, New York, NY, USA, 2025, pp. 1–9. DOI: <https://doi.org/10.1145/3726101.3726102>.
- **Thi-Quynh-Hoa Nguyen**, Duy-Hai Nguyen, Tuan-Long Tran, Tan-Hai Nguyen, Thanh-Trung Dang, and Xuan-Hien Le. Improving Adaptive Learning by Incorporating Question Attributes: A System for Special Needs Education in Vietnam. In: *Proceedings of the 17th International Conference on Knowledge and Systems Engineering (KSE)*, Da Lat, Vietnam, 2025, pp. 1–6. DOI: <https://doi.org/10.1109/KSE68178.2025.11309532>.
- **Thi-Quynh-Hoa Nguyen**, Duy-Hai Nguyen, Tuan-Long Tran, Thi-Trang Tran, and Van-Khoa Le. A Computational Framework for the

Personalized Remediation of Reading Difficulties Using Dynamic Bayesian Networks. In: *Proceedings of the 14th International Symposium on Information and Communication Technology (SoICT 2025)*, Nha Trang, Vietnam, 2025. (**Best Paper Award**)

DOI: https://doi.org/10.1007/978-981-92-2590-3_31

b. Các công trình đã nộp (Đang trong quá trình xét duyệt)

- **Thi-Quynh-Hoa Nguyen**, Nu-Tam-An Nguyen, Thi-Cam-Huong Nguyen, Thanh-Ha Le, and Thi-Duyen Ngo. An automated eye-tracking framework: From semantic gaze mapping to human-in-the-loop optimization. Submitted to *Autism Research*, 2026 (**WoS/SCOPUS Q1**)
- **Thi-Quynh-Hoa Nguyen**, Thi-Duyen Ngo, Duy-Hai Nguyen, Tuan-Long Tran, and Van-Khoa Le. A dual-loop ITS architecture integrating TAD graph backtracking and adaptive format scaffolding: Optimizing literacy intervention. *Submitted to Journal of Computer Assisted Learning*, 2026 (**WoS/SCOPUS Q1**).

KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Kết luận

Luận án đã nghiên cứu và đề xuất các phương pháp khai thác dữ liệu chuyển động mắt nhằm hỗ trợ phân tích hành vi và phát triển các giải pháp công nghệ hỗ trợ trẻ có rối loạn phát triển, tập trung vào rối loạn phổ tự kỉ, rối loạn tính toán và chứng khó đọc. Theo mục tiêu đặt ra, nghiên cứu được triển khai theo chuỗi gồm: xây dựng dữ liệu thực nghiệm; trích xuất đặc trưng, phân tích và tăng cường dữ liệu chuyển động mắt; và khai thác các đặc trưng hành vi để hỗ trợ đánh giá và cá nhân hóa can thiệp.

- Thứ nhất, về dữ liệu thực nghiệm, luận án đã xây dựng hai bộ dữ liệu cho các nhóm trẻ mục tiêu. Bộ dữ liệu RLPTK được thu thập trên 55 trẻ, gồm 28 trẻ RLPTK và 27 trẻ phát triển điển hình, với 660 phiên đo và khoảng 4,6 triệu mẫu điểm nhìn. Bộ dữ liệu RLTT được xây dựng trên 41 trẻ, tích hợp đồng bộ dữ liệu chuyển động mắt, âm thanh và thao tác tương tác, gồm 82 phiên với khoảng 1,93 triệu bản ghi. Các quy trình thiết kế nhiệm vụ, phát triển hệ thống thu nhận dữ liệu, hiệu chuẩn, thu thập, kiểm soát chất lượng và đồng bộ dữ liệu tạo cơ sở cho các phân tích và mô hình tính toán tiếp theo.
- Thứ hai, về xử lý và khai thác dữ liệu chuyển động mắt, luận án đề xuất phương pháp tự động trích xuất đặc trưng thị giác trên kích thích động thông qua kiến trúc kết hợp phát hiện điểm cố định, phân đoạn đối tượng và theo dõi vùng quan tâm. Kiến trúc đạt tốc độ xử lý xấp xỉ 31,5 FPS và cho phép tự động ánh xạ điểm nhìn vào các đối tượng chuyển động. Trên bộ dữ liệu RLPTK, mô hình SVM đánh giá theo cấp độ đối tượng đạt độ chính xác 92,73%, AUC 0,93 và độ nhạy 96,43%,

cho thấy tiềm năng của các đặc trưng thị giác trong hỗ trợ phân loại RLPTK.

Bên cạnh đó, luận án đề xuất phương pháp tăng cường dữ liệu chuyển động mắt dựa trên tối ưu tiến hóa kết hợp chuyển động Brownian phân số. Phương pháp tạo được các quỹ đạo tổng hợp duy trì đặc tính gần với dữ liệu thực và cho kết quả tốt hơn các phương pháp sinh dữ liệu cơ sở trên một số cấu hình TSTR, cho thấy tiềm năng ứng dụng khi dữ liệu chuyển động mắt có quy mô hạn chế.

- Thứ ba, về hỗ trợ can thiệp cá nhân hóa, luận án nghiên cứu và đề xuất các giải pháp theo hai hướng. Đối với trẻ RLPTK, dữ liệu chuyển động mắt được tích hợp vào chu trình can thiệp có sự tham gia của chuyên gia; các đặc trưng thị giác được phân tích qua hệ cơ sở luật để hỗ trợ nhận diện trạng thái hành vi và đưa ra khuyến nghị điều chỉnh. Thực nghiệm trường hợp đơn trên ba trẻ ghi nhận sự thay đổi theo hướng tích cực của các chỉ số hành vi thị giác giữa pha cơ sở và pha can thiệp, cung cấp bằng chứng ban đầu về tính khả thi của cách tiếp cận.

Đối với nhóm trẻ có khó khăn trong học tập, luận án phát triển các nền tảng EzMath và EzRead nhằm hỗ trợ điều chỉnh học liệu, giao diện và hình thức tương tác theo đặc điểm của từng trẻ. Các thử nghiệm bước đầu ghi nhận tính khả dụng và những thay đổi tích cực ở các trường hợp nghiên cứu. Trên cơ sở đó, luận án đề xuất kiến trúc vòng lặp kép kết hợp đồ thị phụ thuộc nhiệm vụ và mô hình truy vết tri thức Bayes nhằm định vị các kỹ năng nền tảng còn thiếu hụt và hỗ trợ lựa chọn nội dung can thiệp; đồng thời đề xuất mô hình DBN mở rộng có khả năng tích hợp đặc trưng chuyển động mắt.

Nhìn chung, đóng góp của luận án nằm ở việc kết nối ba lớp bài toán dữ liệu – xử lý và mô hình hóa hành vi thị giác – hỗ trợ quyết định và can thiệp, qua đó cung cấp cơ sở kỹ thuật cho việc khai thác công nghệ theo dõi chuyển động mắt trong các hệ thống hỗ trợ trẻ rối loạn phát triển. Các kết quả liên quan đến luận án đã được công bố qua 03 bài báo trên tạp chí quốc tế thuộc danh mục WoS/Scopus Q1, 04 công trình tại hội nghị quốc tế, trong đó 01 công trình đạt Best Paper Award, 01 bài báo trên tạp chí trong nước, 01 báo cáo khoa học tại diễn đàn châu Á và 01 đơn đăng ký sáng chế được chấp nhận hợp lệ.

Hạn chế

Mặc dù đạt được các kết quả nêu trên, luận án vẫn còn một số hạn chế. Trước hết, quy mô mẫu còn hạn chế, đặc biệt đối với các thực nghiệm can thiệp trường hợp đơn. Các mô hình phân loại hiện mới được đánh giá trên dữ liệu thu thập trong phạm vi nghiên cứu, chưa được kiểm chứng trên các tập dữ liệu độc lập hoặc tại các cơ sở khác nhau.

Bên cạnh đó, các đặc trưng chuyển động mắt được khai thác chủ yếu là các thông tin không gian – thời gian như vị trí, số lượng và thời lượng điểm cố định, quỹ đạo quan sát. Một số thông tin sinh lý có khả năng phản ánh tải nhận thức, chẳng hạn biến thiên kích thước đồng tử, chưa được khai thác đầy đủ. Các thực nghiệm can thiệp cũng chưa có giai đoạn theo dõi dài hạn để đánh giá khả năng duy trì và chuyển giao kỹ năng.

Định hướng nghiên cứu

Trong thời gian tới, nghiên cứu sẽ mở rộng quy mô và tính đa dạng của mẫu, triển khai kiểm chứng độc lập tại nhiều cơ sở và kiểm soát tốt hơn các yếu tố nền như tuổi, giới tính và năng lực nhận thức nhằm đánh giá khả năng tổng quát hóa của các mô hình.

Một hướng trọng tâm là tích hợp trực tiếp dữ liệu chuyển động mắt vào EzMath và EzRead để hình thành vòng phản hồi khép kín giữa trạng thái hành vi của trẻ và cơ chế điều chỉnh giao diện, học liệu hoặc độ khó nhiệm vụ; đồng thời huấn luyện và kiểm chứng mô hình DBN mở rộng trên dữ liệu thực nghiệm.

Ngoài ra, nghiên cứu sẽ khai thác thêm các nguồn dữ liệu như kích thước đồng tử, âm thanh, hành vi thao tác và các tín hiệu sinh lý khác, kết hợp với các thiết kế thực nghiệm dài hạn. Các giải pháp cũng sẽ được phát triển theo hướng sử dụng thiết bị chi phí thấp, bảo đảm an toàn dữ liệu và thuận lợi cho triển khai trong môi trường giáo dục thực tế.

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

a. Các công trình đã được công bố/đăng tải

1. **Thi-Quynh-Hoa Nguyen**, Thi-Xuan-Thu Le, Thi-Cam-Huong Nguyen, Thanh-Ha Le, and Thi-Duyen Ngo. EzRead: A Mobile-Based Approach for Vietnamese Children with Dyslexia. In: *Proceedings of the 2025 7th Asia Pacific Information Technology Conference (APIT '25)*, Association for Computing Machinery, New York, NY, USA, 2025, pp. 1–9. DOI: <https://doi.org/10.1145/3726101.3726102>.
2. **Thi-Quynh-Hoa Nguyen**, Thi-Kieu-Chinh Nguyen, Thanh-Ha Le, and Thi-Duyen Ngo. Eye-Tracking Technology Applications for Supporting Individuals with Autism Spectrum Disorder: Insights, Challenges, and Opportunities. In: *Disability and Rehabilitation: Assistive Technology*, 2025, 20(8), pp. 2630–2665. DOI: <https://doi.org/10.1080/17483107.2025.2531236>(WoS/SCOPUS Q1, IF: 2.1).
3. **Thi-Quynh-Hoa Nguyen**, Thi-Cam-Huong Nguyen, Thanh-Ha Le, and Thi-Duyen Ngo. EzMath: Developing Digital Methods, a Website-Based Approach to Support Teaching Mathematics to First-Grade Vietnamese Children with Dyscalculia. In: *Journal of Science Natural Science*, 2025, 70(3), pp. 83–94. DOI: <https://doi.org/10.18173/2354-1059.2025-0039>.
4. **Thi-Quynh-Hoa Nguyen**, Duy-Hai Nguyen, Tuan-Long Tran, Tan-Hai Nguyen, Thanh-Trung Dang, and Xuan-Hien Le. Improving Adaptive Learning by Incorporating Question Attributes: A System for Special Needs Education in Vietnam. In: *Proceedings of the 17th International Conference on Knowledge and Systems Engineering (KSE)*, Da Lat, Vietnam, 2025, pp. 1–6. DOI: <https://doi.org/10.1109/KSE68178.2025.11309532>.
5. **Thi-Quynh-Hoa Nguyen**, Duy-Hai Nguyen, Tuan-Long Tran, Thi-Trang Tran, and Van-Khoa Le. A Computational Framework for the

Personalized Remediation of Reading Difficulties Using Dynamic Bayesian Networks. In: *Proceedings of the 14th International Symposium on Information and Communication Technology (SoICT 2025)*, Nha Trang, Vietnam, 2025. (**Best Paper Award**)

DOI: https://doi.org/10.1007/978-981-92-2590-3_8

6. **Thi-Quynh-Hoa Nguyen**, Duy-Hai Nguyen, Thanh-Ha Le, and Thi-Duyen Ngo. An Eye-Tracking System for Extracting and Visualizing Visual Features of Dyscalculia in Children. In: *Proceedings of the 14th International Symposium on Information and Communication Technology (SoICT 2025)*, Nha Trang, Vietnam, 2025. DOI: https://doi.org/10.1007/978-981-92-2590-3_31
7. **Thi-Quynh-Hoa Nguyen**, Tuan-Long Tran, Quang-Huy Trinh, and Thi-Duyen Ngo. Using Technology to Support Children with Dyscalculia in Pre-Primary and Primary School: Insights, Challenges, and Opportunities. In: *Disability and Rehabilitation: Assistive Technology*, 2026, pp. 1–24. DOI: <https://doi.org/10.1080/17483107.2026.2658104> (**WoS/SCOPUS Q1, IF: 2.1**).
8. Đơn đăng ký sáng chế đã được chấp nhận. Số đăng ký: 1-2024-07110. Tên sáng chế: “Phương pháp và hệ thống cung cấp thông tin về sự chú ý thị giác”
9. Báo cáo chuyên đề tại Diễn đàn Châu Á lần thứ nhất về hỗ trợ trẻ rối loạn phổ tự kỷ và trẻ rối loạn phát triển khác. Tên báo cáo: “Ứng dụng công nghệ theo dõi mắt để hiểu đặc điểm nhận thức và tối ưu hoá can thiệp cho trẻ rối loạn phổ tự kỷ”. Địa điểm: Hà Nội, Việt Nam, 06/03/2026.
10. **Thi-Quynh-Hoa Nguyen**, Duy-Hai Nguyen, Gia-Loc Tran, Thanh-Ha Le, and Thi-Duyen Ngo. Bridging the Data Gap: A Genetic Algorithm Approach Using Inverse Python Implementation of Eye Movement Metrics and Visualization Logic. *Applied Soft Computing*, 2026 (**WoS/SCOPUS Q1, IF: 7.8**). DOI: <https://doi.org/10.1016/j.asoc.2026.116328>

b. Các công trình đã nộp (Đang trong quá trình xét duyệt)

11. **Thi-Quynh-Hoa Nguyen**, Duc-Duy Le, Dang-Khoa Ta, Nu-Tam-An Nguyen, Thi-Cam-Huong Nguyen, Thanh-Ha Le, and Thi-Duyen Ngo. A Zero-Shot Segmentation Framework Using Eye-Tracking and the Segment Anything Model: Automated Extraction of Visual Biomarkers in Autism Spectrum Disorder. Submitted to *Engineering Applications of Artificial Intelligence*, 2026 (**WoS/SCOPUS Q1**).

12. **Thi-Quynh-Hoa Nguyen**, Nu-Tam-An Nguyen, Thi-Cam-Huong Nguyen, Thanh-Ha Le, and Thi-Duyen Ngo. An automated eye-tracking framework: From semantic gaze mapping to human-in-the-loop optimization. Submitted to *Autism Research*, 2026 (**WoS/SCOPUS Q1**).
13. **Thi-Quynh-Hoa Nguyen**, Thi-Duyen Ngo, Duy-Hai Nguyen, Tuan-Long Tran, and Van-Khoa Le. A dual-loop ITS architecture integrating TAD graph backtracking and adaptive format scaffolding: Optimizing literacy intervention. Submitted to *Journal of Computer Assisted Learning*, 2026 (**WoS/SCOPUS Q1**).

TÀI LIỆU THAM KHẢO

- [1] A. Poole and L. J. Ball, “Eye tracking in hci and usability research,” in *Encyclopedia of Human Computer Interaction*. IGI Global, 2006, p. 211–219. [Online]. Available: <http://dx.doi.org/10.4018/978-1-59140-562-7.ch034>
- [2] B. T. Carter and S. G. Luke, “Best practices in eye tracking research,” *International Journal of Psychophysiology*, vol. 155, p. 49–62, Sept 2020. [Online]. Available: <http://dx.doi.org/10.1016/j.ijpsycho.2020.05.010>
- [3] J. Sigut and S.-A. Sidha, “Iris center corneal reflection method for gaze tracking using visible light,” *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 2, p. 411–419, Feb 2011. [Online]. Available: <http://dx.doi.org/10.1109/tbme.2010.2087330>
- [4] A. T. Duchowski, *Eye Tracking Methodology: Theory and Practice*. Springer London, 2003. [Online]. Available: <http://dx.doi.org/10.1007/978-1-4471-3750-4>
- [5] J. M. Findlay and I. D. Gilchrist, *Active Vision*. Oxford University Press, Aug. 2003. [Online]. Available: <http://dx.doi.org/10.1093/acprof:oso/9780198524793.001.0001>
- [6] K. Holmqvist, M. Nystrom, R. Andersson, R. Dewhurst, H. Jarodzka, and J. van de Weijer, *Eye Tracking*. London, England: Oxford University Press, Sep. 2011.
- [7] T. J. Anderson and M. R. MacAskill, “Eye movements in patients with neurodegenerative disorders,” *Nature Reviews Neurology*, vol. 9, no. 2, p. 74–85, Jan 2013. [Online]. Available: <http://dx.doi.org/10.1038/nrneurol.2012.273>
- [8] A. Wolf and K. Ueda, “Contribution of eye-tracking to study cognitive impairments among clinical populations,” *Frontiers in Psychology*, vol. 12, June 2021. [Online]. Available: <http://dx.doi.org/10.3389/fpsyg.2021.590986>
- [9] R. Cooper, *Diagnostic and Statistical Manual of Mental Disorders (DSM)*. IMR Press, 2017, vol. 44, no. 8. [Online]. Available: <http://dx.doi.org/10.5771/0943-7444-2017-8-668>

- [10] J. Zagermann, U. Pfeil, and H. Reiterer, “Measuring cognitive load using eye tracking technology in visual computing,” in *Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization*, ser. BELIV ’16. ACM, Oct 2016, p. 78–85. [Online]. Available: <http://dx.doi.org/10.1145/2993901.2993908>
- [11] M. Ekin, K. Krejtz, C. Duarte, A. T. Duchowski, and I. Krejtz, “Prediction of intrinsic and extraneous cognitive load with oculometric and biometric indicators,” *Scientific Reports*, vol. 15, no. 1, Feb 2025. [Online]. Available: <http://dx.doi.org/10.1038/s41598-025-89336-y>
- [12] M. T. KULP and P. P. SCHMIDT, “Visual predictors of reading performance in kindergarten and first grade children,” *Optometry and Vision Science*, vol. 73, no. 4, p. 255–262, Apr 1996. [Online]. Available: <http://dx.doi.org/10.1097/00006324-199604000-00007>
- [13] N. Alsharif, M. H. Al-Adhaileh, M. Al-Yaari, N. Farhah, and Z. I. Khan, “Utilizing deep learning models in an intelligent eye-tracking system for autism spectrum disorder diagnosis,” *Frontiers in Medicine*, vol. 11, July 2024. [Online]. Available: <http://dx.doi.org/10.3389/fmed.2024.1436646>
- [14] L. Rello and M. Ballesteros, “Detecting readers with dyslexia using machine learning with eye tracking measures,” in *Proceedings of the 12th International Web for All Conference*, ser. W4A ’15. ACM, May 2015, p. 1–8. [Online]. Available: <http://dx.doi.org/10.1145/2745555.2746644>
- [15] C. Gillberg, “The essence in child psychiatry: Early symptomatic syndromes eliciting neurodevelopmental clinical examinations,” *Research in Developmental Disabilities*, vol. 31, no. 6, p. 1543–1551, Nov 2010. [Online]. Available: <http://dx.doi.org/10.1016/j.ridd.2010.06.002>
- [16] D. E. Astle, J. Holmes, R. Kievit, and S. E. Gathercole, “Annual research review: The transdiagnostic revolution in neurodevelopmental disorders,” *Journal of Child Psychology and Psychiatry*, vol. 63, no. 4, p. 397–417, July 2021. [Online]. Available: <http://dx.doi.org/10.1111/jcpp.13481>
- [17] G. R. Lyon, S. E. Shaywitz, and B. A. Shaywitz, “A definition of dyslexia,” *Annals of Dyslexia*, vol. 53, no. 1, p. 1–14, Jan 2003. [Online]. Available: <http://dx.doi.org/10.1007/s11881-003-0001-9>
- [18] K. Kucian and M. von Aster, “Developmental dyscalculia,” *European Journal of Pediatrics*, vol. 174, no. 1, p. 1–13, Dec 2014. [Online]. Available: <http://dx.doi.org/10.1007/s00431-014-2455-7>
- [19] M. Turner, “Annotation: Repetitive behaviour in autism: A review of psychological research,” *Journal of Child Psychology and Psychiatry*, vol. 40, no. 6, p. 839–849, Sept 1999. [Online]. Available: <http://dx.doi.org/10.1111/1469-7610.00502>

- [20] B. Butterworth, S. Varma, and D. Laurillard, “Dyscalculia: From brain to education,” *Science*, vol. 332, no. 6033, p. 1049–1053, May 2011. [Online]. Available: <http://dx.doi.org/10.1126/science.1201536>
- [21] K. Landerl, A. Bevan, and B. Butterworth, “Developmental dyscalculia and basic numerical capacities: a study of 8–9-year-old students,” *Cognition*, vol. 93, no. 2, p. 99–125, Sept 2004. [Online]. Available: <http://dx.doi.org/10.1016/j.cognition.2003.11.004>
- [22] F. R. Vellutino, J. M. Fletcher, M. J. Snowling, and D. M. Scanlon, “Specific reading disability (dyslexia): what have we learned in the past four decades?” *Journal of Child Psychology and Psychiatry*, vol. 45, no. 1, p. 2–40, Jan 2004. [Online]. Available: <http://dx.doi.org/10.1046/j.0021-9630.2003.00305.x>
- [23] M. J. Guralnick, “Why early intervention works: A systems perspective,” *Infants & Young Children*, vol. 24, no. 1, p. 6–28, Jan. 2011. [Online]. Available: <http://dx.doi.org/10.1097/iy.0b013e3182002cfe>
- [24] T. Q. H. Nguyen, T. K. C. Nguyen, T. H. Le, and T. D. Ngo, “Eye-tracking technology applications for supporting individuals with autism spectrum disorder: insights, challenges, and opportunities,” *Disability and Rehabilitation: Assistive Technology*, vol. 20, no. 8, p. 2630–2665, July 2025. [Online]. Available: <http://dx.doi.org/10.1080/17483107.2025.2531236>
- [25] E. I. Toki, “Using eye-tracking to assess dyslexia: A systematic review of emerging evidence,” *Education Sciences*, vol. 14, no. 11, p. 1256, Nov. 2024. [Online]. Available: <http://dx.doi.org/10.3390/educsci14111256>
- [26] T. K. C. Nguyen, D. D. Le, T. H. Le, T. C. H. Nguyen, and T. D. Ngo, “The use of eye tracking in supporting individuals with dyslexia: a review,” *Disability and Rehabilitation: Assistive Technology*, vol. 20, no. 5, p. 1183–1198, Dec. 2024. [Online]. Available: <http://dx.doi.org/10.1080/17483107.2024.2437697>
- [27] A. Klin, W. Jones, R. Schultz, F. Volkmar, and D. Cohen, “Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism,” *Archives of General Psychiatry*, vol. 59, no. 9, p. 809, Sept 2002. [Online]. Available: <http://dx.doi.org/10.1001/archpsyc.59.9.809>
- [28] K. Pierce, D. Conant, R. Hazin, R. Stoner, and J. Desmond, “Preference for geometric patterns early in life as a risk factor for autism,” *Archives of General Psychiatry*, vol. 68, no. 1, p. 101, Jan 2011. [Online]. Available: <http://dx.doi.org/10.1001/archgenpsychiatry.2010.113>

- [29] K. Pierce, S. Marinero, R. Hazin, B. McKenna, C. C. Barnes, and A. Malige, “Eye tracking reveals abnormal visual preference for geometric images as an early biomarker of an autism spectrum disorder subtype associated with increased symptom severity,” *Biological Psychiatry*, vol. 79, no. 8, p. 657–666, Apr 2016. [Online]. Available: <http://dx.doi.org/10.1016/j.biopsych.2015.03.032>
- [30] M. Chita-Tegmark, “Social attention in asd: A review and meta-analysis of eye-tracking studies,” *Research in Developmental Disabilities*, vol. 48, p. 79–93, Jan 2016. [Online]. Available: <http://dx.doi.org/10.1016/j.ridd.2015.10.011>
- [31] F. Hutzler and H. Wimmer, “Eye movements of dyslexic children when reading in a regular orthography,” *Brain and Language*, vol. 89, no. 1, p. 235–242, Apr 2004. [Online]. Available: [http://dx.doi.org/10.1016/s0093-934x\(03\)00401-2](http://dx.doi.org/10.1016/s0093-934x(03)00401-2)
- [32] K. Rayner, “Eye movements in reading and information processing: 20 years of research.” *Psychological Bulletin*, vol. 124, no. 3, p. 372–422, 1998. [Online]. Available: <http://dx.doi.org/10.1037/0033-2909.124.3.372>
- [33] J. Hyönä and R. K. Olson, “Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency.” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 21, no. 6, p. 1430–1440, Nov 1995. [Online]. Available: <http://dx.doi.org/10.1037/0278-7393.21.6.1430>
- [34] K. Moeller, S. Neuburger, L. Kaufmann, K. Landerl, and H.-C. Nuerk, “Basic number processing deficits in developmental dyscalculia: Evidence from eye tracking,” *Cognitive Development*, vol. 24, no. 4, p. 371–386, Jan 2009. [Online]. Available: <http://dx.doi.org/10.1016/j.cogdev.2009.09.007>
- [35] S. van Viersen, E. M. Slot, E. H. Kroesbergen, J. E. van’t Noordende, and P. P. M. Leseman, “The added value of eye-tracking in diagnosing dyscalculia: a case study,” *Frontiers in Psychology*, vol. 4, 2013. [Online]. Available: <http://dx.doi.org/10.3389/fpsyg.2013.00679>
- [36] M. Nilsson Benfatto, G. Öqvist Seimyr, J. Ygge, T. Pansell, A. Rydberg, and C. Jacobson, “Screening for dyslexia using eye tracking during reading,” *PLOS ONE*, vol. 11, no. 12, p. e0165508, Dec 2016. [Online]. Available: <http://dx.doi.org/10.1371/journal.pone.0165508>
- [37] T. W. Frazier, E. W. Klingemier, S. Parikh, L. Speer, M. S. Strauss, C. Eng, A. Y. Hardan, and E. A. Youngstrom, “Development and validation of objective and quantitative eye tracking-based measures of autism risk and symptom levels,” *Journal of the American Academy*

- of Child & Adolescent Psychiatry*, vol. 57, no. 11, p. 858–866, Nov 2018. [Online]. Available: <http://dx.doi.org/10.1016/j.jaac.2018.06.023>
- [38] S. Fletcher-Watson and S. Hampton, “The potential of eye-tracking as a sensitive measure of behavioural change in response to intervention,” *Scientific Reports*, vol. 8, no. 1, Oct 2018. [Online]. Available: <http://dx.doi.org/10.1038/s41598-018-32444-9>
- [39] S. Goldstein and J. A. Naglieri, “From assessment to intervention: Strategies for identifying and monitoring treatment progress in asd,” *Autism Spectrum News*, 1 2013. [Online]. Available: <https://autismspectrumnews.org/from-assessment-to-intervention-strategies-for-identifying-and-monitoring-treatment>
- [40] G. A. Holleman, I. T. C. Hooge, C. Kemner, and R. S. Hessels, “The ‘real-world approach’ and its problems: A critique of the term ecological validity,” *Frontiers in Psychology*, vol. 11, Apr 2020. [Online]. Available: <http://dx.doi.org/10.3389/fpsyg.2020.00721>
- [41] E. F. Risko, K. Laidlaw, M. Freeth, T. Foulsham, and A. Kingstone, “Social attention with real versus reel stimuli: toward an empirical approach to concerns about ecological validity,” *Frontiers in Human Neuroscience*, vol. 6, 2012. [Online]. Available: <http://dx.doi.org/10.3389/fnhum.2012.00143>
- [42] K. Krejtz, A. T. Duchowski, A. Niedzielska, C. Biele, and I. Krejtz, “Eye tracking cognitive load using pupil diameter and microsaccades with fixed gaze,” *PLOS ONE*, vol. 13, no. 9, p. e0203629, Sept 2018. [Online]. Available: <http://dx.doi.org/10.1371/journal.pone.0203629>
- [43] P. van der Wel and H. van Steenbergen, “Pupil dilation as an index of effort in cognitive control tasks: A review,” *Psychonomic Bulletin & Review*, vol. 25, no. 6, p. 2005–2015, Feb 2018. [Online]. Available: <http://dx.doi.org/10.3758/s13423-018-1432-y>
- [44] T. D. Parsons, “Virtual reality for enhanced ecological validity and experimental control in the clinical, affective and social neurosciences,” *Frontiers in Human Neuroscience*, vol. 9, Dec. 2015. [Online]. Available: <http://dx.doi.org/10.3389/fnhum.2015.00660>
- [45] J. Sweller, “Cognitive load during problem solving: Effects on learning,” *Cognitive Science*, vol. 12, no. 2, p. 257–285, Apr 1988. [Online]. Available: http://dx.doi.org/10.1207/s15516709cog1202_4
- [46] R. S. Hessels, C. Kemner, C. van den Boomen, and I. T. C. Hooge, “The area-of-interest problem in eyetracking research: A noise-robust solution for face and sparse stimuli,” *Behavior Research Methods*, vol. 48, no. 4, p. 1694–1712, Nov 2015. [Online]. Available: <http://dx.doi.org/10.3758/s13428-015-0676-y>

- [47] I. T. C. Hooge, M. Nyström, D. C. Niehorster, R. Andersson, T. Foulsham, A. Nuthmann, and R. S. Hessels, “The fundamentals of eye tracking part 6: Working with areas of interest,” *Behavior Research Methods*, vol. 58, no. 3, Feb 2026. [Online]. Available: <http://dx.doi.org/10.3758/s13428-025-02937-3>
- [48] R. Pelánek, “Bayesian knowledge tracing, logistic models, and beyond: an overview of learner modeling techniques,” *User Modeling and User-Adapted Interaction*, vol. 27, no. 3-5, p. 313–350, May 2017. [Online]. Available: <http://dx.doi.org/10.1007/s11257-017-9193-2>
- [49] E. B. Delabarre, “A method of recording eye-movements,” *The American Journal of Psychology*, vol. 9, no. 4, p. 572, 1898. [Online]. Available: <http://dx.doi.org/10.2307/1412191>
- [50] E. F. Buchner, “Review of the psychology and pedagogy of reading, with a review of the history of reading and writing and of methods, texts, and hygiene in reading,” *Psychological Bulletin*, vol. 6, no. 4, p. 147–150, Apr 1909. [Online]. Available: <http://dx.doi.org/10.1037/h0066540>
- [51] A. Kennedy, *Book Review: Eye Tracking: A Comprehensive Guide to Methods and Measures*. SAGE Publications, Mar 2016, vol. 69, no. 3. [Online]. Available: <http://dx.doi.org/10.1080/17470218.2015.1098709>
- [52] B. Mahanama, Y. Jayawardana, S. Rengarajan, G. Jayawardena, L. Chukoskie, J. Snider, and S. Jayarathna, “Eye movement and pupil measures: A review,” *Frontiers in Computer Science*, vol. 3, Jan. 2022. [Online]. Available: <http://dx.doi.org/10.3389/fcomp.2021.733531>
- [53] R. S. Hessels and I. T. Hooge, “Eye tracking in developmental cognitive neuroscience – the good, the bad and the ugly,” *Developmental Cognitive Neuroscience*, vol. 40, p. 100710, Dec 2019. [Online]. Available: <http://dx.doi.org/10.1016/j.dcn.2019.100710>
- [54] D. D. Salvucci and J. H. Goldberg, “Identifying fixations and saccades in eye-tracking protocols,” in *Proceedings of the symposium on Eye tracking research & applications - ETRA '00*, ser. ETRA '00. ACM Press, 2000, p. 71–78. [Online]. Available: <http://dx.doi.org/10.1145/355017.355028>
- [55] K. Rayner, “The 35th sir frederick bartlett lecture: Eye movements and attention in reading, scene perception, and visual search,” *Quarterly Journal of Experimental Psychology*, vol. 62, no. 8, p. 1457–1506, Aug. 2009. [Online]. Available: <http://dx.doi.org/10.1080/17470210902816461>
- [56] B. Fischer and H. Weber, “Express saccades and visual attention,” *Behavioral and Brain Sciences*, vol. 16, no. 3, p. 553–567, 1993. [Online]. Available: <http://dx.doi.org/10.1017/s0140525x00031575>

- [57] T. Blascheck, K. Kurzhals, M. Raschke, M. Burch, D. Weiskopf, and T. Ertl, “Visualization of eye tracking data: A taxonomy and survey,” *Computer Graphics Forum*, vol. 36, no. 8, p. 260–284, Feb 2017. [Online]. Available: <http://dx.doi.org/10.1111/cgf.13079>
- [58] D. Noton and L. Stark, “Scanpaths in eye movements during pattern perception,” *Science*, vol. 171, no. 3968, p. 308–311, Jan 1971. [Online]. Available: <http://dx.doi.org/10.1126/science.171.3968.308>
- [59] E. A. D. Lagmay and M. M. T. Rodrigo*, “Enhanced automatic areas of interest (aoi) bounding boxes estimation algorithm for dynamic eye-tracking stimuli,” *APSIPA Transactions on Signal and Information Processing*, vol. 11, no. 1, p. 1–41, July 2022. [Online]. Available: <http://dx.doi.org/10.1561/116.00000026>
- [60] R. K. Wang, K. Kwong, K. Liu, and X.-J. Kong, “New eye tracking metrics system: the value in early diagnosis of autism spectrum disorder,” *Frontiers in Psychiatry*, vol. 15, Dec 2024. [Online]. Available: <http://dx.doi.org/10.3389/fpsy.2024.1518180>
- [61] X.-J. Kong, Z. Wei, B. Sun, Y. Tu, Y. Huang, M. Cheng, S. Yu, G. Wilson, J. Park, Z. Feng, M. Vangel, J. Kong, and G. Wan, “Different eye tracking patterns in autism spectrum disorder in toddler and preschool children,” *Frontiers in Psychiatry*, vol. 13, Jun. 2022. [Online]. Available: <http://dx.doi.org/10.3389/fpsy.2022.899521>
- [62] P. Bonifacci, V. Tobia, A. Sansavini, and A. Guarini, “Eye-movements in a text reading task: A comparison of preterm children, children with dyslexia and typical readers,” *Brain Sciences*, vol. 13, no. 3, p. 425, Feb 2023. [Online]. Available: <http://dx.doi.org/10.3390/brainsci13030425>
- [63] V. del Barrio, *Diagnostic and Statistical Manual of Mental Disorders*. Elsevier, 2017. [Online]. Available: <http://dx.doi.org/10.1016/b978-0-12-809324-5.05530-9>
- [64] J. E. Harrison, S. Weber, R. Jakob, and C. G. Chute, “Icd-11: an international classification of diseases for the twenty-first century,” *BMC Medical Informatics and Decision Making*, vol. 21, no. S6, Nov 2021. [Online]. Available: <http://dx.doi.org/10.1186/s12911-021-01534-6>
- [65] A. Caspi and T. E. Moffitt, “Gene–environment interactions in psychiatry: joining forces with neuroscience,” *Nature Reviews Neuroscience*, vol. 7, no. 7, p. 583–590, July 2006. [Online]. Available: <http://dx.doi.org/10.1038/nrn1925>
- [66] T. Gaugler, L. Klei, S. J. Sanders, C. A. Bodea, A. P. Goldberg, A. B. Lee, M. Mahajan, D. Manaa, Y. Pawitan, J. Reichert, S. Ripke, S. Sandin, P. Sklar, O. Svantesson, A. Reichenberg, C. M. Hultman, B. Devlin, K. Roeder, and J. D. Buxbaum, “Most

- genetic risk for autism resides with common variation,” *Nature Genetics*, vol. 46, no. 8, p. 881–885, July 2014. [Online]. Available: <http://dx.doi.org/10.1038/ng.3039>
- [67] C. Lord, T. S. Brugha, T. Charman, J. Cusack, G. Dumas, T. Frazier, E. J. H. Jones, R. M. Jones, A. Pickles, M. W. State, J. L. Taylor, and J. Veenstra-VanderWeele, “Autism spectrum disorder,” *Nature Reviews Disease Primers*, vol. 6, no. 1, Jan 2020. [Online]. Available: <http://dx.doi.org/10.1038/s41572-019-0138-4>
- [68] M. Frost-Karlsson, M. A. Galazka, C. Gillberg, C. Gillberg, C. Miniscalco, E. Billstedt, N. Hadjikhani, and J. Åsberg Johnels, “Social scene perception in autism spectrum disorder: An eye-tracking and pupillometric study,” *Journal of Clinical and Experimental Neuropsychology*, vol. 41, no. 10, p. 1024–1032, Jul. 2019. [Online]. Available: <http://dx.doi.org/10.1080/13803395.2019.1646214>
- [69] M. B. Al Mourad and E. Bataineh, “Comparative eye-tracking analysis of visual behavior of normal developing and autistic children in response to human faces and emotional expressions,” *Advances in Autism*, vol. 12, no. 1, p. 1–25, Oct. 2025. [Online]. Available: <http://dx.doi.org/10.1108/AIA-09-2024-0061>
- [70] J. M. Moriuchi, A. Klin, and W. Jones, “Mechanisms of diminished attention to eyes in autism,” *American Journal of Psychiatry*, vol. 174, no. 1, p. 26–35, Jan 2017. [Online]. Available: <http://dx.doi.org/10.1176/appi.ajp.2016.15091222>
- [71] T. W. Frazier, M. Strauss, E. W. Klingemier, E. E. Zetzer, A. Y. Hardan, C. Eng, and E. A. Youngstrom, “A meta-analysis of gaze differences to social and nonsocial information between individuals with and without autism,” *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 56, no. 7, p. 546–555, July 2017. [Online]. Available: <http://dx.doi.org/10.1016/j.jaac.2017.05.005>
- [72] B. Noris, J. Nadel, M. Barker, N. Hadjikhani, and A. Billard, “Investigating gaze of children with asd in naturalistic settings,” *PLoS ONE*, vol. 7, no. 9, p. e44144, Sept 2012. [Online]. Available: <http://dx.doi.org/10.1371/journal.pone.0044144>
- [73] S. V. Wass, E. J. H. Jones, T. Gliga, T. J. Smith, T. Charman, M. H. Johnson, S. Baron-Cohen, R. Bedford, P. Bolton, S. Chandler, K. Davies, J. Fernandes, H. Garwood, K. Hudry, H. Maris, G. Pasco, A. Pickles, H. Ribiero, L. Tucker, and A. Volein, “Shorter spontaneous fixation durations in infants with later emerging autism,” *Scientific Reports*, vol. 5, no. 1, Feb. 2015. [Online]. Available: <http://dx.doi.org/10.1038/srep08284>
- [74] B. Butterworth, “Foundational numerical capacities and the origins of dyscalculia,” *Trends in Cognitive Sciences*, vol. 14, no. 12, p. 534–541,

- Dec 2010. [Online]. Available: <http://dx.doi.org/10.1016/j.tics.2010.09.007>
- [75] R. S. Shalev, J. Auerbach, O. Manor, and V. Gross-Tsur, “Developmental dyscalculia: prevalence and prognosis,” *European Child & Adolescent Psychiatry*, vol. 9, no. S2, p. S58–S64, June 2000. [Online]. Available: <http://dx.doi.org/10.1007/s007870070009>
- [76] M. M. Porras, C. A. N. K.-v. Campen, J. J. González-Rosa, F. L. Sánchez-Fernández, and J. I. N. Guzmán, “Eye tracking study in children to assess mental calculation and eye movements,” *Scientific Reports*, vol. 14, no. 1, Aug 2024. [Online]. Available: <http://dx.doi.org/10.1038/s41598-024-69800-x>
- [77] A. R. Strohmaier, K. J. MacKay, A. Obersteiner, and K. M. Reiss, “Eye-tracking methodology in mathematics education research: A systematic literature review,” *Educational Studies in Mathematics*, vol. 104, no. 2, p. 147–200, June 2020. [Online]. Available: <http://dx.doi.org/10.1007/s10649-020-09948-1>
- [78] M. Wolf and P. G. Bowers, “The double-deficit hypothesis for the developmental dyslexias,” *Journal of Educational Psychology*, vol. 91, no. 3, p. 415–438, Sept 1999. [Online]. Available: <http://dx.doi.org/10.1037/0022-0663.91.3.415>
- [79] R. L. Peterson and B. F. Pennington, “Developmental dyslexia,” *The Lancet*, vol. 379, no. 9830, p. 1997–2007, May 2012. [Online]. Available: [http://dx.doi.org/10.1016/S0140-6736\(12\)60198-6](http://dx.doi.org/10.1016/S0140-6736(12)60198-6)
- [80] M. P. Bucci, D. Brémond-Gignac, and Z. Kapoula, “Poor binocular coordination of saccades in dyslexic children,” *Graefes’ Archive for Clinical and Experimental Ophthalmology*, vol. 246, no. 3, p. 417–428, Nov 2007. [Online]. Available: <http://dx.doi.org/10.1007/s00417-007-0723-1>
- [81] B. Kitchenham, S. Charters *et al.*, “Guidelines for performing systematic literature reviews in software engineering,” 2007. [Online]. Available: https://legacyfileshare.elsevier.com/promis_misc/525444systematicreviewsguide.pdf
- [82] P. Schleifer and K. Landerl, “Subitizing and counting in typical and atypical development,” *Developmental Science*, vol. 14, no. 2, p. 280–291, Feb. 2011. [Online]. Available: <http://dx.doi.org/10.1111/j.1467-7687.2010.00976.x>
- [83] N. Derakshan and E. H. Koster, “Processing efficiency in anxiety: Evidence from eye-movements during visual search,” *Behaviour Research and Therapy*, vol. 48, no. 12, p. 1180–1185, Dec. 2010. [Online]. Available: <http://dx.doi.org/10.1016/j.brat.2010.08.009>

- [84] M. R. Kanhirakadavath and M. S. M. Chandran, “Investigation of eye-tracking scan path as a biomarker for autism screening using machine learning algorithms,” *Diagnostics*, vol. 12, no. 2, p. 518, Feb 2022. [Online]. Available: <http://dx.doi.org/10.3390/diagnostics12020518>
- [85] I. A. Ahmed, E. M. Senan, T. H. Rassem, M. A. H. Ali, H. S. A. Shatnawi, S. M. Alwazer, and M. Alshahrani, “Eye tracking-based diagnosis and early detection of autism spectrum disorder using machine learning and deep learning techniques,” *Electronics*, vol. 11, no. 4, p. 530, Feb. 2022. [Online]. Available: <http://dx.doi.org/10.3390/electronics11040530>
- [86] D. Pitta-Pantazi, E. Demosthenous, M. Schindler, A. J. Lilienthal, and C. Christou, “Structure sense in students’ quantity comparison and repeating pattern extension tasks: an eye-tracking study with first graders,” *Educational Studies in Mathematics*, vol. 118, no. 3, p. 339–357, Jan. 2024. [Online]. Available: <http://dx.doi.org/10.1007/s10649-023-10290-5>
- [87] L. Baumanns, D. Pitta-Pantazi, C. Christou, A. J. Lilienthal, A. L. Simon, and M. Schindler, “Adaptive strategy use in pattern-recognition of first graders with and without risk of developing mathematical difficulties: An eye-tracking study,” in *Proceedings of the 46th Conference of the International Group for the Psychology of Mathematics Education (PME46)*, M. Ayalon, B. Koichu, R. Leikin, L. Rubel, and M. Tabach, Eds., vol. 2. PME, 2023, pp. 75–82. [Online]. Available: <https://www.igpme.org/wp-content/uploads/2024/06/PME46-2023-Haifa.zip>
- [88] B. Nerušil, J. Polec, J. Škunda, and J. Kačur, “Eye tracking based dyslexia detection using a holistic approach,” *Scientific Reports*, vol. 11, no. 1, Aug. 2021. [Online]. Available: <http://dx.doi.org/10.1038/s41598-021-95275-1>
- [89] E. Pellicano and M. Stears, “Bridging autism, science and society: moving toward an ethically informed approach to autism research,” *Autism Research*, vol. 4, no. 4, p. 271–282, May 2011. [Online]. Available: <http://dx.doi.org/10.1002/aur.201>
- [90] D. W. Sosnowski, C. O. Stough, M. J. Weiss, T. Cessna, A. Casale, A. Foran, M. Erwinski, J. Wilson, S. A. Farber, and M. A. Farber, “Brief report: A novel digital therapeutic that combines applied behavior analysis with gaze-contingent eye tracking to improve emotion recognition in children with autism spectrum disorder,” *Journal of Autism and Developmental Disorders*, vol. 52, no. 5, p. 2357–2366, Jun. 2021. [Online]. Available: <http://dx.doi.org/10.1007/s10803-021-05101-w>
- [91] S. D’Mello, A. Olney, C. Williams, and P. Hays, “Gaze tutor: A gaze-reactive intelligent tutoring system,” *International Journal of*

- Human-Computer Studies*, vol. 70, no. 5, p. 377–398, May 2012. [Online]. Available: <http://dx.doi.org/10.1016/j.ijhcs.2012.01.004>
- [92] G. Dawson, E. J. Jones, K. Merkle, K. Venema, R. Lowy, S. Faja, D. Kamara, M. Murias, J. Greenson, J. Winter, M. Smith, S. J. Rogers, and S. J. Webb, “Early behavioral intervention is associated with normalized brain activity in young children with autism,” *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 51, no. 11, p. 1150–1159, Nov 2012. [Online]. Available: <http://dx.doi.org/10.1016/j.jaac.2012.08.018>
- [93] W. Jones and A. Klin, “Attention to eyes is present but in decline in 2–6-month-old infants later diagnosed with autism,” *Nature*, vol. 504, no. 7480, p. 427–431, Nov 2013. [Online]. Available: <http://dx.doi.org/10.1038/nature12715>
- [94] D. D. Price, D. G. Finniss, and F. Benedetti, “A comprehensive review of the placebo effect: Recent advances and current thought,” *Annual Review of Psychology*, vol. 59, no. 1, p. 565–590, Jan. 2008. [Online]. Available: <http://dx.doi.org/10.1146/annurev.psych.59.113006.095941>
- [95] H. Liu, “Personalized learning support system for special education: a real-time feedback mechanism based on deep reinforcement learning,” *Frontiers in Psychology*, vol. 16, Nov 2025. [Online]. Available: <http://dx.doi.org/10.3389/fpsyg.2025.1658698>
- [96] T. Käser, A. G. Busetto, B. Solenthaler, G.-M. Baschera, J. Kohn, K. Kucian, M. von Aster, and M. Gross, *Modelling and Optimizing Mathematics Learning in Children*. Springer Science and Business Media LLC, Nov. 2013, vol. 23, no. 1–4. [Online]. Available: <http://dx.doi.org/10.1007/s40593-013-0003-7>
- [97] A. Gupta, N. Sendhilnathan, J. Hartcher-O’Brien, E. Pezent, H. Benko, and T. R. Jonker, “Investigating eyes-away mid-air typing in virtual reality using squeeze haptics-based postural reinforcement,” in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, ser. CHI ’23. ACM, Apr 2023, p. 1–11. [Online]. Available: <http://dx.doi.org/10.1145/3544548.3581467>
- [98] L. Rello and R. Baeza-Yates, “How to present more readable text for people with dyslexia,” *Universal Access in the Information Society*, vol. 16, no. 1, p. 29–49, Nov 2015. [Online]. Available: <http://dx.doi.org/10.1007/s10209-015-0438-8>
- [99] M. H. Schneps, C. Chen, M. Pomplun, J. Wang, A. D. Crosby, and K. Kent, “Pushing the speed of assistive technologies for reading,” *Mind, Brain, and Education*, vol. 13, no. 1, p. 14–29, Dec 2018. [Online]. Available: <http://dx.doi.org/10.1111/mbe.12180>

- [100] L. Rello and J. P. Bigham, “Good background colors for readers: A study of people with and without dyslexia,” in *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*, ser. ASSETS '17. ACM, Oct 2017, p. 72–80. [Online]. Available: <http://dx.doi.org/10.1145/3132525.3132546>
- [101] L. Rello, G. Kanvinde, and R. Baeza-Yates, “Layout guidelines for web text and a web service to improve accessibility for dyslexics,” p. 1–9, Apr 2012. [Online]. Available: <http://dx.doi.org/10.1145/2207016.2207048>
- [102] M. H. Schneps, J. M. Thomson, C. Chen, G. Sonnert, and M. Pomplun, “E-readers are more effective than paper for some with dyslexia,” *PLoS ONE*, vol. 8, no. 9, p. e75634, Sept 2013. [Online]. Available: <http://dx.doi.org/10.1371/journal.pone.0075634>
- [103] J. N. Constantino, S. Kennon-McGill, C. Weichselbaum, N. Marrus, A. Haider, A. L. Glowinski, S. Gillespie, C. Klaiman, A. Klin, and W. Jones, “Infant viewing of social scenes is under genetic control and is atypical in autism,” *Nature*, vol. 547, no. 7663, p. 340–344, July 2017. [Online]. Available: <http://dx.doi.org/10.1038/nature22999>
- [104] R. S. Hessels, R. Andersson, I. T. C. Hooge, M. Nyström, and C. Kemner, “Consequences of eye color, positioning, and head movement for eye-tracking data quality in infant research,” *Infancy*, vol. 20, no. 6, p. 601–633, June 2015. [Online]. Available: <http://dx.doi.org/10.1111/inf.12093>
- [105] N. Valliappan, N. Dai, E. Steinberg, J. He, K. Rogers, V. Ramachandran, P. Xu, M. Shojaeizadeh, L. Guo, K. Kohlhoff, and V. Navalpakkam, “Accelerating eye movement research via accurate and affordable smartphone eye tracking,” *Nature Communications*, vol. 11, no. 1, Sept 2020. [Online]. Available: <http://dx.doi.org/10.1038/s41467-020-18360-5>
- [106] A. M. Mastergeorge, C. Kahathuduwa, and J. Blume, “Eye-tracking in infants and young children at risk for autism spectrum disorder: A systematic review of visual stimuli in experimental paradigms,” *Journal of Autism and Developmental Disorders*, vol. 51, no. 8, p. 2578–2599, Oct 2020. [Online]. Available: <http://dx.doi.org/10.1007/s10803-020-04731-w>
- [107] R. J. Cook, B. M. Dickens, and M. F. Fathalla, “World medical association declaration of helsinki: Ethical principles for medical research involving human subjects,” p. 428–432, Apr 2003. [Online]. Available: <http://dx.doi.org/10.1093/acprof:oso/9780199241323.003.0025>
- [108] M. A. Cascio, J. A. Weiss, and E. Racine, “Person-oriented research ethics to address the needs of participants on the autism spectrum,”

- Ethics & Human Research*, vol. 42, no. 5, p. 2–16, Sept 2020. [Online]. Available: <http://dx.doi.org/10.1002/eahr.500064>
- [109] C. Kothe, S. Y. Shirazi, T. Stenner, D. Medine, C. Boulay, M. I. Grivich, F. Artoni, T. Mullen, A. Delorme, and S. Makeig, “The lab streaming layer for synchronized multimodal recording,” *Imaging Neuroscience*, vol. 3, 2025. [Online]. Available: <http://dx.doi.org/10.1162/IMAG.a.136>
- [110] N. Nowack, J. Wittsiepe, M. Kasper-Sonnenberg, M. Wilhelm, and A. Schölmerich, “Influence of low-level prenatal exposure to pcdd/fs and pcbs on empathizing, systemizing and autistic traits: Results from the duisburg birth cohort study,” *PLOS ONE*, vol. 10, no. 6, p. e0129906, June 2015. [Online]. Available: <http://dx.doi.org/10.1371/journal.pone.0129906>
- [111] K. Friston, “The free-energy principle: a rough guide to the brain?” *Trends in Cognitive Sciences*, vol. 13, no. 7, p. 293–301, July 2009. [Online]. Available: <http://dx.doi.org/10.1016/j.tics.2009.04.005>
- [112] M. Zorzi, C. Barbiero, A. Facoetti, I. Lonciari, M. Carrozzi, M. Montico, L. Bravar, F. George, C. Pech-Georgel, and J. C. Ziegler, “Extra-large letter spacing improves reading in dyslexia,” *Proceedings of the National Academy of Sciences*, vol. 109, no. 28, p. 11455–11459, June 2012. [Online]. Available: <http://dx.doi.org/10.1073/pnas.1205566109>
- [113] U. M. Schaller, M. Biscaldi, A. Burkhardt, C. Fleischhaker, M. Herbert, A. Isringhausen, L. Tebartz van Elst, and R. Rauh, “Ados-eye-tracking: The archimedean point of view and its absence in autism spectrum conditions,” *Frontiers in Psychology*, vol. 12, Mar. 2021. [Online]. Available: <http://dx.doi.org/10.3389/fpsyg.2021.584537>
- [114] U. Bronfenbrenner, “Toward an experimental ecology of human development.” *American Psychologist*, vol. 32, no. 7, p. 513–531, 1977. [Online]. Available: <http://dx.doi.org/10.1037/0003-066X.32.7.513>
- [115] E. A. Holmes, C. Creswell, and T. G. O’Connor, “Intrusive imagery of events of september 11 measure,” 2007. [Online]. Available: <http://dx.doi.org/10.1037/t15018-000>
- [116] A. T. Duchowski, *Eye Tracking Methodology*. Springer International Publishing, 2017. [Online]. Available: <http://dx.doi.org/10.1007/978-3-319-57883-5>
- [117] R. Hersh and S. Dehaene, *The Number Sense: How the Mind Creates Mathematics*. Informa UK Limited, Dec 1998, vol. 105, no. 10. [Online]. Available: <http://dx.doi.org/10.2307/2589308>

- [118] S. Dehaene, “The neural basis of the weber–fechner law: a logarithmic mental number line,” *Trends in Cognitive Sciences*, vol. 7, no. 4, p. 145–147, Apr 2003. [Online]. Available: [http://dx.doi.org/10.1016/s1364-6613\(03\)00055-x](http://dx.doi.org/10.1016/s1364-6613(03)00055-x)
- [119] W. B. Bilker, J. A. Hansen, C. M. Brensinger, J. Richard, R. E. Gur, and R. C. Gur, “Development of abbreviated nine-item forms of the raven’s standard progressive matrices test,” *Assessment*, vol. 19, no. 3, p. 354–369, May 2012. [Online]. Available: <http://dx.doi.org/10.1177/1073191112446655>
- [120] R. S. Shalev, O. Manor, B. Kerem, M. Ayali, N. Badichi, Y. Friedlander, and V. Gross-Tsur, “Developmental dyscalculia is a familial learning disability,” *Journal of Learning Disabilities*, vol. 34, no. 1, p. 59–65, Jan 2001. [Online]. Available: <http://dx.doi.org/10.1177/002221940103400105>
- [121] L. Kaufmann and M. v. Aster, “The diagnosis and management of dyscalculia,” *Deutsches Ärzteblatt international*, Nov 2012. [Online]. Available: <http://dx.doi.org/10.3238/arztebl.2012.0767>
- [122] R. H. Bradley and R. F. Corwyn, “Socioeconomic status and child development,” *Annual Review of Psychology*, vol. 53, no. 1, p. 371–399, Feb 2002. [Online]. Available: <http://dx.doi.org/10.1146/annurev.psych.53.100901.135233>
- [123] S. Kunjan, T. S. Grummett, K. J. Pope, D. M. W. Powers, S. P. Fitzgibbon, T. Bastiampillai, M. Battersby, and T. W. Lewis, *The Necessity of Leave One Subject Out (LOSO) Cross Validation for EEG Disease Diagnosis*. Springer International Publishing, 2021, p. 558–567. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-86993-9_50
- [124] A. T. Duchowski, *Eye Tracking Methodology*. Springer International Publishing, 2017. [Online]. Available: <http://dx.doi.org/10.1007/978-3-319-57883-5>
- [125] V. Krassanakis, V. Filippakopoulou, and B. Nakos, “Eyemmv toolbox: An eye movement post-analysis tool based on a two-step spatial dispersion threshold for fixation identification,” *Journal of Eye Movement Research*, vol. 7, no. 1, Feb 2014. [Online]. Available: <http://dx.doi.org/10.16910/jemr.7.1.1>
- [126] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, P. Dollár, and R. Girshick, “Segment anything,” in *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*. IEEE, Oct 2023, p. 3992–4003. [Online]. Available: <http://dx.doi.org/10.1109/iccv51070.2023.00371>

- [127] B. D. Lucas and T. Kanade, “An iterative image registration technique with an application to stereo vision,” in *Proceedings of the 7th international joint conference on Artificial intelligence*, vol. 2, 1981, pp. 674–679. [Online]. Available: <https://hal.science/hal-03697340/>
- [128] A. Borji and L. Itti, “State-of-the-art in visual attention modeling,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 1, p. 185–207, Jan. 2013. [Online]. Available: <http://dx.doi.org/10.1109/tpami.2012.89>
- [129] C. E. Shannon, “A mathematical theory of communication,” *Bell System Technical Journal*, vol. 27, no. 3, p. 379–423, July 1948. [Online]. Available: <http://dx.doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- [130] B. Sekachev, N. Manovich, and A. Zhavoronkov, “Computer vision annotation tool,” 2019. [Online]. Available: <https://zenodo.org/record/3497106>
- [131] F. Milletari, N. Navab, and S.-A. Ahmadi, “V-net: Fully convolutional neural networks for volumetric medical image segmentation,” in *2016 Fourth International Conference on 3D Vision (3DV)*. IEEE, Oct 2016, p. 565–571. [Online]. Available: <http://dx.doi.org/10.1109/3dv.2016.79>
- [132] K. He, G. Gkioxari, P. Dollar, and R. Girshick, “Mask r-cnn,” in *2017 IEEE International Conference on Computer Vision (ICCV)*. IEEE, Oct 2017, p. 2980–2988. [Online]. Available: <http://dx.doi.org/10.1109/iccv.2017.322>
- [133] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, “Attention is all you need,” *Advances in neural information processing systems*, vol. 30, 2017. [Online]. Available: <https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html>
- [134] S. Baker and I. Matthews, “Lucas-kanade 20 years on: A unifying framework,” *International Journal of Computer Vision*, vol. 56, no. 3, p. 221–255, Feb. 2004. [Online]. Available: <http://dx.doi.org/10.1023/b:visi.0000011205.11775.fd>
- [135] *New Metropolitan Perspectives: Knowledge Dynamics and Innovation-driven Policies Towards Urban and Regional Transition Volume 2*. Springer International Publishing, 2021. [Online]. Available: <http://dx.doi.org/10.1007/978-3-030-48279-4>
- [136] T. Q. H. Nguyen, D. D. Le, D. K. Ta, N. T. A. Nguyen, T. C. H. Nguyen, T. H. Le, and T. D. Ngo, “A zero-shot segmentation framework using eye-tracking and the segment anything model: Automated extraction of visual biomarkers in autism spectrum disorder,” June 2026, submitted to Engineering Applications of Artificial Intelligence.

- [137] *Pattern Recognition and Machine Learning*. Springer US, 1971. [Online]. Available: <http://dx.doi.org/10.1007/978-1-4615-7566-5>
- [138] B. Efron, “Bootstrap methods: Another look at the jackknife.” Institute of Mathematical Statistics, Jan 1979, vol. 7, no. 1. [Online]. Available: <http://dx.doi.org/10.1214/aos/1176344552>
- [139] S. Sidney, *NONPARAMETRIC STATISTICS FOR THE BEHAVIORAL SCIENCES*. Ovid Technologies (Wolters Kluwer Health), July 1957, vol. 125, no. 3. [Online]. Available: <http://dx.doi.org/10.1097/00005053-195707000-00032>
- [140] H. B. Mann and D. R. Whitney, “On a test of whether one of two random variables is stochastically larger than the other,” *The Annals of Mathematical Statistics*, vol. 18, no. 1, p. 50–60, Mar 1947. [Online]. Available: <http://dx.doi.org/10.1214/aoms/1177730491>
- [141] C. Spearman, “The proof and measurement of association between two things,” *The American Journal of Psychology*, vol. 15, no. 1, p. 72, Jan 1904. [Online]. Available: <http://dx.doi.org/10.2307/1412159>
- [142] R. I. Parker, K. J. Vannest, J. L. Davis, and S. B. Sauber, “Combining nonoverlap and trend for single-case research: Tau-u,” *Behavior Therapy*, vol. 42, no. 2, p. 284–299, June 2011. [Online]. Available: <http://dx.doi.org/10.1016/j.beth.2010.08.006>
- [143] R. A. Fisher, “Statistical methods for research workers,” in *Breakthroughs in Statistics*. Springer New York, 1992, p. 66–70. [Online]. Available: http://dx.doi.org/10.1007/978-1-4612-4380-9_6
- [144] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, “Gene selection for cancer classification using support vector machines,” *Machine Learning*, vol. 46, no. 1-3, p. 389–422, Jan. 2002. [Online]. Available: <http://dx.doi.org/10.1023/A:1012487302797>
- [145] U. von Luxburg and B. Schoelkopf, “Statistical learning theory: Models, concepts, and results,” 2008. [Online]. Available: <https://arxiv.org/abs/0810.4752>
- [146] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, *Human Activity Recognition on Smartphones Using a Multiclass Hardware-Friendly Support Vector Machine*. Springer Berlin Heidelberg, 2012, p. 216–223. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-35395-6_30
- [147] M. Alcañiz, I. A. Chicchi-Giglioli, L. A. Carrasco-Ribelles, J. Marín-Morales, M. E. Minissi, G. Teruel-García, M. Sirera, and L. Abad, “Eye gaze as a biomarker in the recognition of autism spectrum disorder using virtual reality and machine learning: A proof of concept for diagnosis,” *Autism Research*, vol. 15, no. 1, p. 131–145, Nov 2021. [Online]. Available: <http://dx.doi.org/10.1002/aur.2636>

- [148] J. S. Oliveira, F. O. Franco, M. C. Revers, A. F. Silva, J. Portolese, H. Brentani, A. Machado-Lima, and F. L. S. Nunes, “Computer-aided autism diagnosis based on visual attention models using eye tracking,” *Scientific Reports*, vol. 11, no. 1, May 2021. [Online]. Available: <http://dx.doi.org/10.1038/s41598-021-89023-8>
- [149] R. Carette, M. Elbattah, F. Cilia, G. Dequen, J.-L. Guérin, and J. Bosche, “Learning to predict autism spectrum disorder based on the visual patterns of eye-tracking scanpaths,” in *Proceedings of the 12th International Joint Conference on Biomedical Engineering Systems and Technologies*. SCITEPRESS - Science and Technology Publications, 2019, p. 103–112. [Online]. Available: <http://dx.doi.org/10.5220/0007402601030112>
- [150] B. B. Mandelbrot and J. W. Van Ness, “Fractional brownian motions, fractional noises and applications,” *SIAM Review*, vol. 10, no. 4, p. 422–437, Oct. 1968. [Online]. Available: <http://dx.doi.org/10.1137/1010093>
- [151] R. Koposov, S. Fossum, T. Frodl, Ø. Nytrø, B. Leventhal, A. Sourander, S. Quaglini, M. Molteni, M. de la Iglesia Vayá, H.-U. Prokosch, N. Barbarini, M. P. Milham, F. X. Castellanos, and N. Skokauskas, “Clinical decision support systems in child and adolescent psychiatry: a systematic review,” *European Child & Adolescent Psychiatry*, vol. 26, no. 11, p. 1309–1317, Apr 2017. [Online]. Available: <http://dx.doi.org/10.1007/s00787-017-0992-0>
- [152] E. Bolhuis, J. Voogt, and K. Schildkamp, “The development of data use, data skills, and positive attitude towards data use in a data team intervention for teacher educators,” *Studies in Educational Evaluation*, vol. 60, p. 99–108, Mar 2019. [Online]. Available: <http://dx.doi.org/10.1016/j.stueduc.2018.12.002>
- [153] A. T. Corbett and J. R. Anderson, “Knowledge tracing: Modeling the acquisition of procedural knowledge,” *User Modelling and User-Adapted Interaction*, vol. 4, no. 4, p. 253–278, 1995. [Online]. Available: <http://dx.doi.org/10.1007/BF01099821>
- [154] F. Paas, J. E. Tuovinen, H. Tabbers, and P. W. M. Van Gerven, “Cognitive load measurement as a means to advance cognitive load theory,” *Educational Psychologist*, vol. 38, no. 1, p. 63–71, Jan 2003. [Online]. Available: http://dx.doi.org/10.1207/s15326985ep3801_8
- [155] H.-Y. Suen and K.-E. Hung, “An interpretable closed-loop intelligent tutoring system for multimodal affective feedback in asynchronous presentation training,” 2026. [Online]. Available: <https://arxiv.org/abs/2605.17468>
- [156] I. Karabey Aksakalli, T. Çelik, A. B. Can, and B. Tekinerdogan, “Deployment and communication patterns in microservice

- architectures: A systematic literature review,” *Journal of Systems and Software*, vol. 180, p. 111014, Oct. 2021. [Online]. Available: <http://dx.doi.org/10.1016/j.jss.2021.111014>
- [157] A. M. Hofmeister and J. M. Ferrara, “Expert systems and special education,” *Exceptional Children*, vol. 53, no. 3, p. 235–239, Nov 1986. [Online]. Available: <http://dx.doi.org/10.1177/001440298605300306>
- [158] O. Bulut, J. Shin, S. N. Yildirim-Erbasli, G. Gorgun, and Z. A. Pardos, “An introduction to bayesian knowledge tracing with pybkt,” *Psych*, vol. 5, no. 3, p. 770–786, July 2023. [Online]. Available: <http://dx.doi.org/10.3390/psych5030050>
- [159] S. Malhotra, G. Rajender, M. S. Bhatia, and T. B. Singh, “Effects of picture exchange communication system on communication and behavioral anomalies in autism,” *Indian Journal of Psychological Medicine*, vol. 32, no. 2, p. 141–143, July 2010. [Online]. Available: <http://dx.doi.org/10.4103/0253-7176.78513>
- [160] G. Mesibov, M. Howley, and S. Naftel, *Accessing the Curriculum for Learners with Autism Spectrum Disorders*. Routledge, July 2015. [Online]. Available: <http://dx.doi.org/10.4324/9781315683782>
- [161] J. N. T. Usry, *Validation of the Assessment of Basic Language and Learning Skills-Revised for students with autism spectrum disorder using an expert review panel*. Liberty University, 2015. [Online]. Available: <https://digitalcommons.liberty.edu/doctoral/1056/>
- [162] K. P. Freeman, J. R. Cook, and E. H. Hooijberg, “Introduction to statistical quality control,” p. 733–739, Apr 2021. [Online]. Available: <http://dx.doi.org/10.2460/javma.258.7.733>
- [163] S. Rakap, “Effect sizes as result interpretation aids in single-subject experimental research: description and application of four nonoverlap methods,” *British Journal of Special Education*, vol. 42, no. 1, p. 11–33, Feb. 2015. [Online]. Available: <http://dx.doi.org/10.1111/1467-8578.12091>
- [164] D. Szucs, A. Devine, F. Soltesz, A. Nobes, and F. Gabriel, “Developmental dyscalculia is related to visuo-spatial memory and inhibition impairment,” *Cortex*, vol. 49, no. 10, p. 2674–2688, Nov. 2013. [Online]. Available: <http://dx.doi.org/10.1016/j.cortex.2013.06.007>
- [165] D. M. Levi, “Crowding—an essential bottleneck for object recognition: A mini-review,” *Vision Research*, vol. 48, no. 5, p. 635–654, Feb 2008. [Online]. Available: <http://dx.doi.org/10.1016/j.visres.2007.12.009>
- [166] D. Wood, J. S. Bruner, and G. Ross, “The role of tutoring in problem solving,” *Journal of Child Psychology and Psychiatry*,

- vol. 17, no. 2, p. 89–100, Apr. 1976. [Online]. Available: <http://dx.doi.org/10.1111/j.1469-7610.1976.tb00381.x>
- [167] L. S. VYGOTSKY, *Mind in Society: Development of Higher Psychological Processes*, M. Cole, V. Jolm-Steiner, S. Scribner, and E. Souberman, Eds. Harvard University Press, Oct. 1980. [Online]. Available: <http://dx.doi.org/10.2307/j.ctvjf9vz4>
- [168] T. Q. H. Nguyen, D. H. Nguyen, T. L. Tran, V. K. Le, and T. D. Ngo, “A dual-loop its architecture integrating tad graph backtracking and adaptive format scaffolding: Optimizing literacy intervention,” June 2026, submitted to Journal of Computer Assisted Learning.
- [169] T. Q. H. Nguyen, D. H. Nguyen, T. L. Tran, T. T. Tran, and V. K. Le, “A computational framework for the personalized remediation of reading difficulties using dynamic bayesian networks,” 2025, accepted/Presented at Proceedings of the 14th International Symposium on Information and Communication Technology (SoICT 2025), Nha Trang, Vietnam.
- [170] T. Q. H. Nguyen, D. H. Nguyen, T. L. Tran, T. H. Nguyen, T. T. Dang, and X. H. Le, “Improving adaptive learning by incorporating question attributes: A system for special needs education in vietnam,” in *2025 17th International Conference on Knowledge and System Engineering (KSE)*. IEEE, Nov 2025, p. 1–6. [Online]. Available: <http://dx.doi.org/10.1109/kse68178.2025.11309532>
- [171] R. S. J. d. Baker, A. T. Corbett, and V. Aleven, *More Accurate Student Modeling through Contextual Estimation of Slip and Guess Probabilities in Bayesian Knowledge Tracing*. Springer Berlin Heidelberg, 2008, p. 406–415. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-69132-7_44
- [172] P. C. M. Molenaar, “A dynamic factor model for the analysis of multivariate time series,” *Psychometrika*, vol. 50, no. 2, p. 181–202, June 1985. [Online]. Available: <http://dx.doi.org/10.1007/BF02294246>
- [173] J. E. Beck and K.-m. Chang, *Identifiability: A Fundamental Problem of Student Modeling*. Springer Berlin Heidelberg, p. 137–146. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-73078-1_17
- [174] C. Piech, J. Spencer, J. Huang, S. Ganguli, M. Sahami, L. Guibas, and J. Sohl-Dickstein, “Deep knowledge tracing,” 2015. [Online]. Available: <https://arxiv.org/abs/1506.05908>
- [175] T. Hastie, R. Tibshirani, and J. Friedman, “The elements of statistical learning,” 2009. [Online]. Available: <http://dx.doi.org/10.1007/978-0-387-84858-7>

-
- [176] Y. Zhang, I. Kuzborskij, J. D. Lee, C. Leng, and F. Liu, “Dag-math: Graph-of-thought guided mathematical reasoning in llms,” 2025. [Online]. Available: <https://arxiv.org/abs/2510.19842>
- [177] G. S. Maddala and L.-F. Lee, “Recursive systems containing qualitative endogeneous variables: A reply,” *Econometrica*, vol. 48, no. 3, p. 765, Apr 1980. [Online]. Available: <http://dx.doi.org/10.2307/1913135>