

DỰ ĐOÁN MỨC ĐỘ PHÁ HOẠI CỤC BỘ CỦA TẤM BÊ TÔNG CỐT SỢI CHỊU TẢI TRỌNG VÀ ĐẬP SỬ DỤNG THUẬT TOÁN MÁY HỌC

Lê Đại Nhân^a, Thái Đức Kiên^b, Doãn Quốc Hoàn^c, Nguyễn Đăng Nguyên^a, Phạm Thái Hoàn^{a,*}

^a*Khoa Xây dựng dân dụng và công nghiệp, Trường Đại học Xây dựng Hà Nội, 55 đường Giải Phóng, quận Hai Bà Trưng, Hà Nội, Việt Nam*

^b*Khoa Xây dựng, Đại học Vinh, 182 Lê Duẩn, Vinh, Nghệ An, Việt Nam*

^c*Khoa Kỹ thuật Xây dựng và Môi trường, Đại học Sejong, Seoul, Hàn Quốc*

Nhận ngày 12/01/2022, Sửa xong 23/04/2022, Chấp nhận đăng 06/5/2022

Tóm tắt

Bài báo giới thiệu và ứng dụng thuật toán phân loại Support Vector Machine (SVM) và thuật toán tối ưu hóa Bayesian (Bayesian optimization - BO), những thuật toán máy học (Machine Learning - ML) mạnh để dự đoán mức độ phá hoại cục bộ của tấm bê tông cốt sợi chịu tải trọng và đập gây ra bởi đầu đạn. Bộ dữ liệu kết quả thu được từ 176 thí nghiệm tấm bê tông cốt sợi chịu tải trọng và đập của đầu đạn bao gồm 15 thông số đầu vào được sử dụng để huấn luyện các mô hình máy học nhằm đưa ra các kết quả dự đoán về mức độ phá hoại cục bộ của tấm. Do các dữ liệu thí nghiệm thu thập được có sự phân tán về các thông số đầu vào và mất cân bằng về các dạng phá hoại cục bộ, một kỹ thuật lấy thêm mẫu BorderlineSMOTE (Synthetic Minority Over-Sampling Technique - SMOTE) cũng được sử dụng nhằm mục đích tạo ra bộ dữ liệu phù hợp để mô hình huấn luyện (được gọi là BO-SVM) đạt kết quả dự đoán tốt. Kết quả thu được cho thấy mô hình đề xuất có thể ứng dụng để dự đoán mức độ phá hoại cục bộ của tấm bê tông cốt sợi chịu tải trọng và đập bởi đầu đạn với độ chính xác chấp nhận được và hiệu quả cao hơn các mô hình ML thông thường khác.

Từ khóa: thuật toán phân loại SVM; thuật toán tối ưu hóa BO; kỹ thuật lấy thêm mẫu BorderlineSMOTE; máy học; bê tông cốt sợi; tải trọng và đập.

PREDICTION OF LOCAL DAMAGE OF FIBER REINFORCED CONCRETE UNDER IMPACT LOADING USING MACHINE LEARNING ALGORITHMS

Abstract

This study introduces and applies Support Vector Machine (SVM) classification and Bayesian optimization (BO), the powerful machine learning (ML) algorithms to predict the local damage of FRC slab subjected to the impact load induced by a warhead. The results from 176 experiments of FRC slabs subjected to warhead impact loads including 15 input parameters are used as dataset to train the ML models to produce predictive results on the local damage of FRC slabs. Due to the dispersion of the collected experimental data in terms of input parameters and imbalance in local damage modes, the BorderlineSMOTE (Synthetic Minority Over-Sampling Technique - SMOTE) sampling technique is also used to create a suitable dataset for the training model (called BO-SVM) to achieve good prediction results. The obtained results showed that the proposed BO-SVM model can be applied to predict the local damage of FRC slabs subjected to impact loads by warheads with acceptable accuracy and higher efficiency than the other conventional and common ML models.

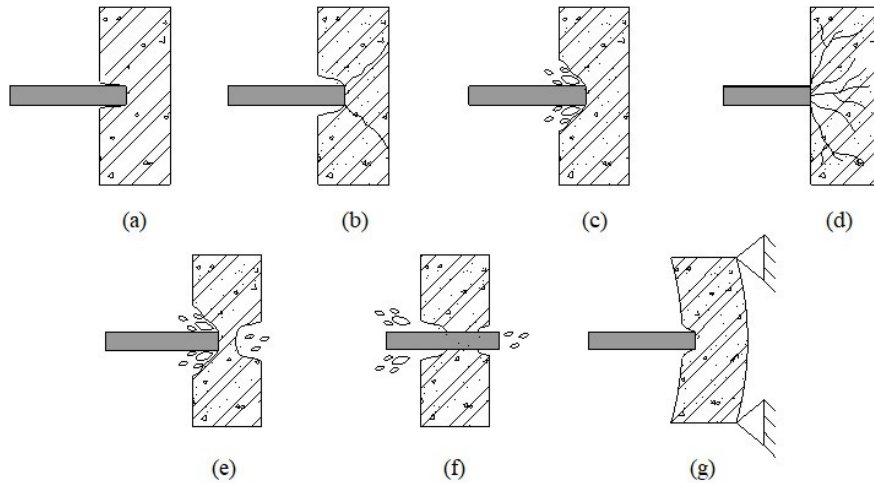
Keywords: support vector machine; Bayesian optimization; synthetic minority over-sampling technique; machine learning; fiber reinforced concrete; impact loading.

[https://doi.org/10.31814/stce.huce\(nuce\)2022-16\(2V\)-03](https://doi.org/10.31814/stce.huce(nuce)2022-16(2V)-03) © 2022 Trường Đại học Xây dựng Hà Nội (ĐHXDHN)

*Tác giả đại diện. Địa chỉ e-mail: hoanpt@huce.edu.vn (Hoàn, P. T.)

1. Giới thiệu

Bê tông nói chung và bê tông cốt thép, cốt sợi nói riêng đã được sử dụng rộng rãi trong nhiều năm bởi các kỹ sư quân sự và dân dụng trong thiết kế và xây dựng các kết cấu bảo vệ để chống va đập và tải trọng nổ. Dưới tác động của các tải trọng va đập (impact) như sự va chạm của đầu đạn, các kết cấu bê tông thường bị phá hoại cục bộ.



Hình 1. Các dạng phá hoại của tấm bê tông dưới tác động va đập của đầu đạn: (a) thâm nhập, (b) nứt hình nón, (c) bong tách, (d) nứt hướng tâm, (e) vỡ, (f) xuyên thủng, và (g) phá hoại tổng thể

Các nghiên cứu thực nghiệm trên rất nhiều mẫu [1–3] chỉ ra rằng có thể xảy ra bảy dạng phá hoại có thể xảy ra với các mục tiêu dưới tác động của đầu đạn, bao gồm: (i) thâm nhập (Penetration), đầu đạn tạo thành một đường hầm vào mục tiêu (Hình 1(a), chiều dài của đường hầm được gọi là độ sâu thâm nhập); (ii) xuất hiện vết nứt hình nón và đạn cắm vào mục tiêu (Cone cracking and plugging), hình thành một vết nứt giống như hình nón dưới đạn và viên đạn cắm vào tấm bê tông (Hình 1(b)); (iii) bong tách (Spalling), vật liệu của tấm bê tông ở mặt trước bị bong ra (Hình 1(c)); (iv) nứt hướng tâm (Radial cracking), các vết nứt tỏa ra từ điểm tác động và xuất hiện trên mặt trước hoặc sau của tấm bê tông hoặc cả hai khi các vết nứt phát triển qua độ dày của mục tiêu (Hình 1(d)); (v) vỡ (Scabbing), các mảnh vỡ bị vỡ và tách ra khỏi mặt sau của tấm bê tông (Hình 1(e)); (vi) xuyên thủng (Perforation), viên đạn hoàn toàn xuyên qua mục tiêu và có hoặc không có vận tốc dư sau khi xuyên qua (Hình 1(f)); và (vii) phá hoại tổng thể của kết cấu (Overall structural failure), tấm bê tông bị phá hoại uốn và chịu cắt tổng thể (Hình 1(g)).

Vật liệu bê tông cốt sợi đã và đang được sử dụng rộng rãi trong xây dựng bởi vì những tính ưu việt của nó trong việc tăng cường độ dẻo và khả năng phân tán năng lượng. Đặc biệt, với khả năng hấp thụ năng lượng rất tốt, bê tông cốt sợi đã và đang được sử dụng phổ biến trong các kết cấu chịu các tải trọng cực hạn như tải trọng nổ và tải trọng va đập do tên lửa [4, 5]. Một trong những nghiên cứu sớm nhất liên quan đến khả năng chịu lực của tấm bê tông cốt sợi đã được thực hiện bởi Dancygier and Yankelevsky [6] bằng thực nghiệm cho thấy cốt sợi tăng cường đáng kể khả năng chịu kháng xuyên của tấm bê tông, đồng thời các vết nứt được phân tán thành nhiều vết nứt nhỏ và giảm thiểu phạm vi phá hoại của tấm. Từ đó, rất nhiều nghiên cứu thực nghiệm đã được thực hiện nhằm đánh giá khả năng chịu kháng xuyên của tấm bê tông cốt sợi với nhiều kích thước cấu kiện, kích thước và vận tốc đầu

đạn, cường độ bê tông và hàm lượng cốt sợi khác nhau [7–10]. Bên cạnh những nỗ lực trong nghiên cứu thực nghiệm, nhằm phục vụ cho công tác thực hành thiết kế kết cấu công trình, một số tác giả đã cố gắng đề xuất một số công thức thực nghiệm để xác định chiều sâu xuyên thủng cũng như tính toán chiều dày tấm tường khi chịu một tải trọng va đập cụ thể [11–13]. Mặc dù đã có nhiều biểu thức thực nghiệm đã được đề xuất, nhưng những biểu thức này chỉ áp dụng phù hợp cho kết cấu bê tông cốt thép thông thường. Mặt khác, mỗi biểu thức cũng chỉ áp dụng được trong một phạm vi nhỏ các tham số được sử dụng trong các thí nghiệm tương ứng. Để tính toán thiết kế kết cấu bê tông cốt sợi, Almusallam và các cộng sự của mình [14, 15] đã đề xuất một số biểu thức thực nghiệm dựa trên các biểu thức đã đề xuất bởi Hội đồng Nghiên cứu Phòng thủ Quốc gia (NDRC) của Hoa Kỳ, dành cho bê tông cốt thép thông thường [14]. Tuy nhiên, các biểu thức này chỉ có thể áp dụng được cho kết cấu bê tông cốt sợi phối hợp giữa cốt sợi thép và cốt sợi polymer trong một phạm vi hữu hạn hàm lượng cốt sợi, và tất yếu không thể sử dụng cho tất cả các trường hợp thiết kế kết cấu trong công trình xây dựng.

Ngày nay, với sự phát triển mạnh mẽ của khoa học máy tính, các kỹ thuật máy học (machine learning-ML) được ứng dụng rộng rãi trong mọi lĩnh vực của cả khoa học và cuộc sống. Các kỹ thuật ML có thể hiểu một cách đơn giản là các phương pháp cho phép xây dựng mô hình toán học phức tạp có độ chính xác rất cao nhằm thể hiện mối quan hệ giữa thông số đầu vào và thông số đầu ra của một tập dữ liệu cho trước. Với quan điểm này, ML được nhiều nhà khoa học hiện nay sử dụng để xác định ứng xử của kết cấu công trình nói chung [16, 17] và mức độ phá hoại của kết cấu bê tông cốt thép, cốt sợi nói riêng chịu các loại tải trọng cực đoan như tải trọng nổ và va đập [18–21]. Các nghiên cứu này cho thấy rằng nếu như chúng ta có một bộ dữ liệu đủ lớn thì các kỹ thuật ML hoàn toàn có thể được sử dụng để xây dựng nên các mô hình hiệu quả để ước lượng ứng xử của công trình nhằm thay thế hoặc hỗ trợ trong việc chính xác hóa các công thức thiết kế truyền thống. Một số thuật toán ML mạnh hiện nay có thể kể đến như là: Gradient Tree Boosting (GTB), học sâu (Deep Learning) (DL), Support Vector Machine (SVM), rừng ngẫu nhiên (Random Forest) (RF), cây quyết định (Decision Tree) (DT), v.v. Hiệu quả của ứng dụng thuật toán ML trong thiết kế công trình phụ thuộc nhiều yếu tố trong đó việc lựa chọn tham số hệ thống phù hợp cho từng lớp bài toán cụ thể là rất quan trọng. Mỗi lớp bài toán sẽ cần được thực hiện, phân tích để tìm ra cách sử dụng các tham số sao cho mô hình huấn luyện đạt hiệu quả cao.

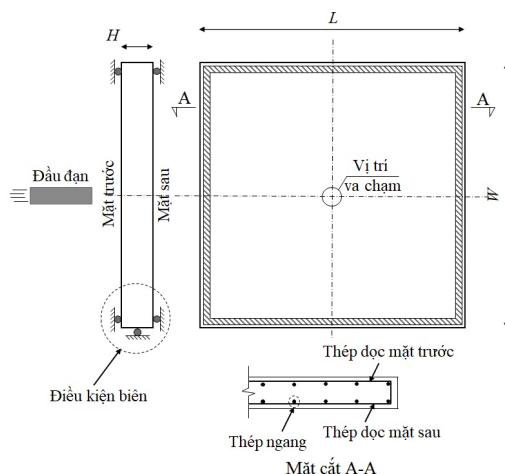
Trên cơ sở đó, trong nghiên cứu này, tác giả trình bày chi tiết cách xây dựng mô hình dự đoán mức độ phá hoại cục bộ của tấm bê tông cốt sợi chịu tải trọng va đập sử dụng kết hợp thuật toán phân loại Support Vector Machine (SVM) và thuật toán tối ưu hóa Bayesian (Bayesian optimization – BO). Bộ dữ liệu kết quả thu được từ 176 thí nghiệm tấm bê tông cốt sợi chịu tải trọng va đập của đầu đạn bao gồm 15 thông số đầu vào được sử dụng để huấn luyện các mô hình máy học nhằm đưa ra các kết quả dự đoán về mức độ phá hoại cục bộ của tấm. Do các dữ liệu thí nghiệm thu thập được có sự phân tán về các thông số đầu vào và mất cân bằng về các dạng phá hoại cục bộ, một kỹ thuật lấy thêm mẫu BorderlineSMOTE (Synthetic Minority Over-Sampling Technique - SMOTE) cũng được sử dụng nhằm mục đích tạo ra bộ dữ liệu phù hợp để mô hình huấn luyện (được gọi là BO-SVM) đạt kết quả dự đoán tốt. Kết quả thu được cho thấy mô hình đề xuất có thể ứng dụng để dự đoán mức độ phá hoại cục bộ của tấm bê tông cốt sợi chịu tải trọng va đập bởi đầu đạn với độ chính xác chấp nhận được và hiệu quả cao hơn các thuật toán ML thông thường khác như mô hình cơ bản SVM trong cả trường hợp có và không sử dụng kỹ thuật lấy thêm mẫu BorderlineSMOTE, mô hình cây quyết định (Decision tree - DT) và mạng nhiều lớp (Multi-layer Perceptron – MLP). Đây là các mô hình được áp dụng phổ biến và được chứng minh có khả năng phân loại tốt trong các nghiên cứu thuộc lĩnh vực xây dựng và môi trường [22–24].

2. Tập dữ liệu thí nghiệm

Hình 2 minh họa thí nghiệm tấm bê tông cốt sợi dưới tác động va đập của đầu đạn điển hình trong đó đầu đạn được bắn vuông góc vào tấm bê tông cốt sợi, đây là góc đặc trưng phổ biến được thực hiện trong nhiều nghiên cứu và thường diễn ra trong thực tế. Trong thí nghiệm được mô tả ở Hình 2, đầu đạn có thể có các đặc trưng như hình dạng, tính chất, khối lượng, vận tốc va đập khác nhau. Khi chịu tác động va đập, tấm bê tông cốt sợi có thể bị phá hoại tổng thể hoặc cục bộ, tuy nhiên đối với vận tốc lớn như vận tốc trong khoảng thông thường của đầu đạn thì dạng phá hoại thường là cục bộ như đã được trình bày ở trên. Trong thực hành thiết kế kết cấu, chỉ 4 trạng thái được kể đến bao gồm 3 dạng phá hoại phổ quát (có thể bao hàm các dạng phá hoại gần tương tự) là xuyên thủng (*perforation*), vỡ (*scabbing*), thâm nhập (*penetration*) và trạng thái là không bị phá hoại (*no damage*). Tiêu chuẩn thiết kế an toàn cho nhà máy hạt nhân của Mỹ (ACI 349-01) [25] cũng quy định rằng các trạng thái giới hạn về phá hoại của kết cấu là xuyên thủng (*perforation*), vỡ (*scabbing*), trong đó xuyên thủng là tình huống xấu nhất. Do bộ dữ liệu thí nghiệm thu thập được không đề cập đến trạng thái không phá hoại, 3 dạng phá hoại còn lại là thủng (*perforation*), vỡ (*scabbing*), và thâm nhập (*penetration*) sẽ được coi là các tiêu chí mức độ phá hoại cục bộ của tấm bê tông cốt sợi dưới tác động va đập của đầu đạn.

Tập dữ liệu được thu thập từ các thí nghiệm tấm bê tông cốt sợi (FRC) chịu tải trọng va đập của đầu đạn trong các nghiên cứu trước đây đã được công bố [6, 8, 10, 14, 15, 26, 27]. Tập dữ liệu bao gồm 176 thí nghiệm, được phân chia làm 3 nhóm phá hoại theo từng cấp độ: *penetration* (51 thí nghiệm), *scabbing* (33 thí nghiệm), và *perforation* (92 thí nghiệm). Các thông số đầu vào của thí nghiệm được sử dụng trong việc huấn luyện mô hình máy học bao gồm 15 thông số, mô tả các thành phần chính, ảnh hưởng đến sự phá hoại của tấm FRC, bao gồm: kích thước tấm, điều kiện biên, đặc tính và thành phần cốt thép, cốt sợi sử dụng, cường độ của bê tông, và các thông số của đầu đạn. Cụ thể, kích thước của các tấm FRC trong khoảng từ $300 \times 170 \text{ mm}^2$ đến $800 \times 800 \text{ mm}^2$, với bề dày tấm giao động từ 50 cho đến 200 mm. Cường độ bê tông từ 28,8 MPa đến 237 MPa. Thành phần cốt thép bao gồm tỷ lệ lượng thép dọc sử dụng trong lưới thép trước, sau, cốt thép ngang và đặc biệt là tỷ lệ sợi đã dùng trong bê tông. Cụ thể, hàm lượng cốt thép mặt trước (mặt va chạm) và mặt sau của tấm dao động từ 0 đến 4,75%, hàm lượng cốt thép ngang dao động từ 0 đến 0,57%, trong khi hàm lượng cốt sợi từ 0 đến 3,0%. Các thông số đầu đạn bao gồm: loại đầu đạn (cứng hoặc mềm), đường kính, trọng lượng và loại mũi (nhọn, bầu, ...). Đầu đạn có đường kính thay đổi trong khoảng từ 7,92 mm đến 76 mm, khối lượng từ 0,008 kg đến 1,575 kg, trong khi đó vận tốc bắn được thí nghiệm nằm trong khoảng từ 84 m/s đến 728 m/s. Điều kiện biên bao gồm: loại tấm, được phân loại theo phương chịu lực (một phương: one-way hoặc hai phương: two-way) và loại gối đỡ là ngàm hoặc khớp. Mô tả chi tiết của các thông số đầu vào được thể hiện trong Bảng 1.

Các thí nghiệm được thu thập đều với các đầu đạn được bắn theo phương thẳng góc với tấm bê tông và nhắm vào vị trí chính giữa của tấm. Do đó, các thông số như phương bắn hoặc vị trí tiếp xúc của đầu đạn trên tấm bê tông cốt sợi không được kể đến trong tập dữ liệu huấn luyện. Các thông



Hình 2. Minh họa thí nghiệm tấm bê tông cốt sợi chịu tải va đập của đầu đạn

Nhân, L. Đ., và cs. / Tạp chí Khoa học Công nghệ Xây dựng
 Bảng 1. Mô tả thông số đầu vào và đầu ra của tập dữ liệu

		Mô tả	Ký hiệu
Đầu vào	Kích thước tấm FRC	Chiều dài tấm FRC	L
		Chiều rộng tấm FRC	W
		Độ dày	H
	Điều kiện biên	Loại tấm: 1 phương (1), hai phương (2)	$Ptype$
		Điều kiện biên: Gối ở 4 góc của tấm (0), Ngàm ở 4 cạnh của tấm (1)	$BCtype$
	Thành phần cốt thép	Tỷ lệ cốt dọc phía trước	ρ_{LF}
		Tỷ lệ cốt dọc phía trước	ρ_{LR}
		Tỷ lệ cốt ngang	ρ_S
		Tỷ lệ cốt sợi (fiber)	ρ_f
	Cường độ bê tông	Cường độ chịu nén của bê-tông	f_{ck}
Thông số đầu đạn	Loại đạn : mềm (0), cứng (1)	$Mtype$	
	Đường kính đầu đạn	d	
	Trọng lượng đầu đạn	M	
	Loại mũi đạn (hệ số): Flat (0,72), Blunt (0,84), Spherical (1,00), Hollow/flat (1,03), Bi-conic (1,05), Ogi-val (1,10), Sharp (1,14)	N	
	Vận tốc bắn	V_0	
Đầu ra	Dạng phá hoại	Dạng phá hoại: Penetration (1,0), Scabbing (2,0), Perforation (3,0)	Y

số đầu vào được kể trên bao gồm cả các thông số dạng số (numerical) và thông số dạng phân loại (categorical). Do đó, trước khi đưa vào huấn luyện, các thông số cần được điều chỉnh để việc huấn luyện được nhanh chóng và đạt hiệu quả cao. Cụ thể, tất cả các thông số dạng số được chuẩn hóa và nằm trong khoảng [0, 1]. Việc này đảm bảo các thông số đầu vào có trọng số là tương đương nhau khi huấn luyện. Các thông số dạng phân loại được chuyển thành một véc-tơ dạng số, gồm các thành phần 0 và 1. Bảng 2 minh họa số liệu chuẩn hóa của 3 thí nghiệm ứng với 3 dạng phá hoại khác nhau thu thập từ nghiên cứu của Riedel và cs. [27].

Bảng 2. Minh họa chuẩn hóa dữ liệu thí nghiệm dùng trong mô hình huấn luyện

Thông số	Đặc trưng	Đơn vị	Mẫu thí nghiệm		
			1	2	3
L	X1	mm	700	700	700
W	X2	mm	700	400	700
H	X3	mm	100	100	100
$Ptype$	X4		2	2	2
$BCtype$	X5		0	0	0
ρ_{LF}	X6	%	4,75	4,75	4,75
ρ_{LR}	X7	%	4,75	4,75	4,75
ρ_S	X8	%	0	0	0
ρ_f	X9	%	1,00	1,00	1,00
f_{ck}	X10	MPa	182,8	172,1	186,0
$Mtype$	X11		0	0	0
d	X12	mm	76	76	76
M	X13	kg	1,564	1,569	1,570
N	X14		0,72	0,72	0,72
V_0	X15	m/s	194,7	258,7	320,0
Y	Y		Penetration	Scabbing	Perforation

3. Các thuật toán

3.1. Thuật toán Support Vector Machine (SVM)

Thuật toán Support vector machine (SVM) là một thuật toán phi xác suất được sử dụng rộng rãi cho các vấn đề phân loại [28]. Mục tiêu của thuật toán SVM là tìm một siêu mặt phẳng trong không gian N chiều (N số đặc điểm). Siêu mặt phẳng là ranh giới quyết định giúp phân loại rõ ràng các điểm dữ liệu. Một siêu mặt phẳng không tổng quát có thể được tạo ra nếu các điểm dữ liệu không được tách biệt tuyến tính. Để giải quyết vấn đề này, một công thức tính vùng biên mềm (soft margin) được giới thiệu trong thuật toán SVM cho phép một số trường hợp bị phân loại sai từ đó giúp tránh hiện tượng quá khớp (overfitting). Thuật toán SVM có nhược điểm là kém hiệu quả trên các bộ dữ liệu bị nhiễu với các lớp chồng chéo nhau. Trong nghiên cứu này, thuật toán SVM sẽ được sử dụng để phân loại các dạng phá hoại cục bộ của tấm bê tông cốt sợi dưới tải tác động của tải trọng va đập gây ra bởi đầu đạn. Tham số tổng quát C , loại hàm “kernel” (kernel function) và tham số “kernel” γ có ảnh hưởng lớn đến kết quả dự đoán của mô hình SVM. Để tìm ra các giá trị tối ưu của các siêu tham số (hyperparameters) này cho mô hình SVM, phương pháp tối ưu hóa (Bayesian optimization – BO) sẽ được sử dụng.

3.2. Phương pháp tối ưu Bayesian (Bayesian optimization-BO)

Phương pháp tối ưu hóa Bayesian (Bayesian optimization-BO) [29] là một phương pháp nổi tiếng trong thực tế lĩnh vực học máy, chủ yếu được sử dụng để điều chỉnh các tham số của mô hình học máy. Cần lưu ý rằng có rất nhiều phương pháp tối ưu hóa được sử dụng phổ biến hiện nay [30], tuy nhiên thuật toán tối ưu BO được xem là phù hợp với bài toán có thông số đầu vào phân bố trong khoảng như

trong nghiên cứu này. BO được biết đến như một cách tiếp cận dựa trên mô hình liên tục để tìm kiếm các cực trị tổng thể của một hàm $f(x)$ chưa biết trên một số miền giới hạn \mathcal{X} :

$$x^* = \arg \max_{x \in \mathcal{X}} f(x) \quad (1)$$

BO hoạt động chủ yếu bằng cách xây dựng một mô hình thay thế xác suất của $f(x)$, trong đó có một phân phối các điểm mô phỏng các tính chất của $f(x)$. Sau đó, những giá trị có ở trong mô hình đại diện được sử dụng để tạo ra hàm thu thập $a(x)$. Điểm kiểm tra x_t tiếp theo được xác định bằng cách tối ưu hàm $a(x)$: $x_t = \arg \max_x a(x)$. Hàm thu thập (Acquisition Function) $a(x)$ là hàm dùng để tìm hiểu không gian của tham số trong quá trình tối ưu Bayesian, hàm sử dụng giá trị trung bình và phương sai dự đoán được tạo ra bởi quy trình Gaussian (Gaussian process - GP). Sau đó, hiệu quả của hàm $f(x)$ được đánh giá bằng các tham số x_t vừa cập nhật. Quá trình này diễn ra cho đến khi đạt được các thông số tốt nhất. Trong nghiên cứu này, quy trình Gaussian được sử dụng để tạo mô hình đại diện vì sự linh hoạt cũng như khả năng phân phối xác suất trước của nó. GP được xác định bởi tính chất rằng bất kỳ tập hợp hữu hạn nào của N điểm $\{x_i \in \mathcal{X}\}_{i=1}^N$ đều tạo ra sự phân bố Gaussian đa biến trên \mathbb{R}^N . Hàm phân bố chuẩn được đặc trưng bởi giá trị trung bình $\mu(x)$ và phương sai $\sigma^2(x)$. Cần lưu ý rằng GP cũng có các tham số riêng như các tham số ban đầu (Base_Estimator), hàm thu thập (Acquisition Function), hay cách tối ưu hàm thu thập (Acquisition optimizer). Về hàm thu thập, nói chung, hàm dựa trên quan sát trước đó và các tham số GP. Có nhiều sự lựa chọn phổ biến khác nhau của hàm thu thập như: xác suất cải thiện (probability of improvement), cải thiện dự kiến (expected improvement - EI), giới hạn trên của độ tin cậy (upper confidence bounds - UCB), vv. Nghiên cứu này tập trung vào hàm EI bởi vì nó cho hiệu quả tốt trong việc giảm thiểu các vấn đề và không yêu cầu điều chỉnh các thông số của riêng nó. Hàm EI có thể được thể hiện như sau:

$$a(x) = EI(x) = \begin{cases} (\mu(x) - f(\hat{x}))\Phi(Z) + \sigma(x)\phi(Z), & \text{nếu } \sigma(x) > 0 \\ 0, & \text{nếu } \sigma(x) = 0 \end{cases} \quad (2)$$

Với $Z = \frac{\mu(x) - f(\hat{x})}{\sigma(x)}$

trong đó \hat{x} là tham số tốt nhất đã ghi nhận được, $\Phi(\cdot)$ và $\phi(\cdot)$ là các hàm phân bố tích lũy và hàm mật độ xác suất của phân bố tiêu chuẩn Gaussian. Hàm EI bao gồm 2 phần khi $\sigma(x) > 0$ thì có thể hiểu là sự đánh đổi giữa việc tập trung tìm kiếm vùng tối ưu đã biết và khai phá những vùng chưa biết của hàm mục tiêu.

3.3. Kỹ thuật lấy thêm mẫu BorderlineSMOTE

Do sự mất cân bằng của bộ dữ liệu, kỹ thuật lấy thêm mẫu khá phổ biến BorderlineSMOTE được sử dụng [31]. Kỹ thuật này hoạt động bằng cách tạo dữ liệu dựa trên kỹ thuật lấy thêm mẫu thiểu số tổng hợp (Synthetic Minority Over-Sampling Technique - SMOTE) [32]. Cần nhấn mạnh rằng, kỹ thuật lấy thêm mẫu SMOTE có nhiều biến thể khác nhau như SMOTE_ENN, SMOTE_TomekLinks, ADASYN, SafelevelSMOTE, tuy nhiên kỹ thuật Borderline SMOTE được sử dụng trong nghiên cứu này do đặc thù của bộ dữ liệu có tính phân bố không đều và chồng lấn lên nhau [33]. Vì các trường hợp gần đường biên giới (nơi các dữ liệu của một lớp gần với các lớp khác) dễ bị phân loại sai hơn các trường hợp cách xa đường biên, những trường hợp này có trọng số cao hơn và cần phải được dành nhiều sự chú ý hơn. Theo đó, lớp thiểu số gần ranh giới sẽ được lấy thêm mẫu dựa trên cơ chế lấy mẫu dữ liệu của SMOTE. Trong nghiên cứu này, bộ dữ liệu được chia thành năm phần bằng cách sử dụng quy trình xác nhận chéo k -lần (cross-validation k -fold). Trong số đó, một phần sẽ được dùng

để kiểm tra và các phần còn lại được sử dụng để đào tạo mô hình. Để ngăn chặn vấn đề lạc quan quá mức (overoptimistic) [33], BorderlineSMOTE sẽ được sử dụng bên trong vòng xác nhận chéo (cross-validation). Tất cả các lớp được lấy thêm mẫu sẽ không bao gồm lớp đa số, ở đây là lớp thủng (Perforation). Đặc biệt, lớp thâm nhập (Penetration) và lớp bị vỡ (Scabbing) đã được lấy thêm mẫu lên đến 73 trường hợp lần lượt từ 40 trường hợp lớp thâm nhập (Penetration), và 26 trường hợp lớp bị vỡ (Scabbing).

3.4. Mô hình đề xuất BO-SVM

Trong nghiên cứu này, mức độ phá hoại cục bộ của tấm FRC được dự đoán chủ yếu dựa trên mô hình SVM. Ba siêu tham số chính bao gồm tham số tổng quát C , loại hàm “kernel” (kernel function) và tham số “kernel” γ thường có ảnh hưởng đáng kể đến hiệu suất của mô hình SVM, do đó tối ưu hóa Bayesian đã được sử dụng để xác định giá trị tốt nhất của các tham số này và sau đó được sử dụng để xây dựng mô hình cuối cùng BO-SVM nhằm dự đoán mức độ phá hoại cục bộ của tấm FRC do tác động va đập của đầu đạn gây nên. Bốn công thức phổ biến của hàm trung tâm bao gồm: Linear, Polynomial, Gaussian RBF, Sigmoid. Công thức toán học của các hàm được thể hiện như sau:

$$\text{Linear: } K(a, b) = a^T b \quad (3)$$

$$\text{Polynomial: } K(a, b) = (\gamma a^T b + r)^d \quad (4)$$

$$\text{Gaussian RBF: } K(a, b) = e^{(-\gamma \|a-b\|^2)} \quad (5)$$

$$\text{Sigmoid: } K(a, b) = \tanh(\gamma a^T b + r) \quad (6)$$

Sơ đồ thuật toán của mô hình BO-SVM đề xuất bao gồm 6 bước được minh họa ở Hình 3.

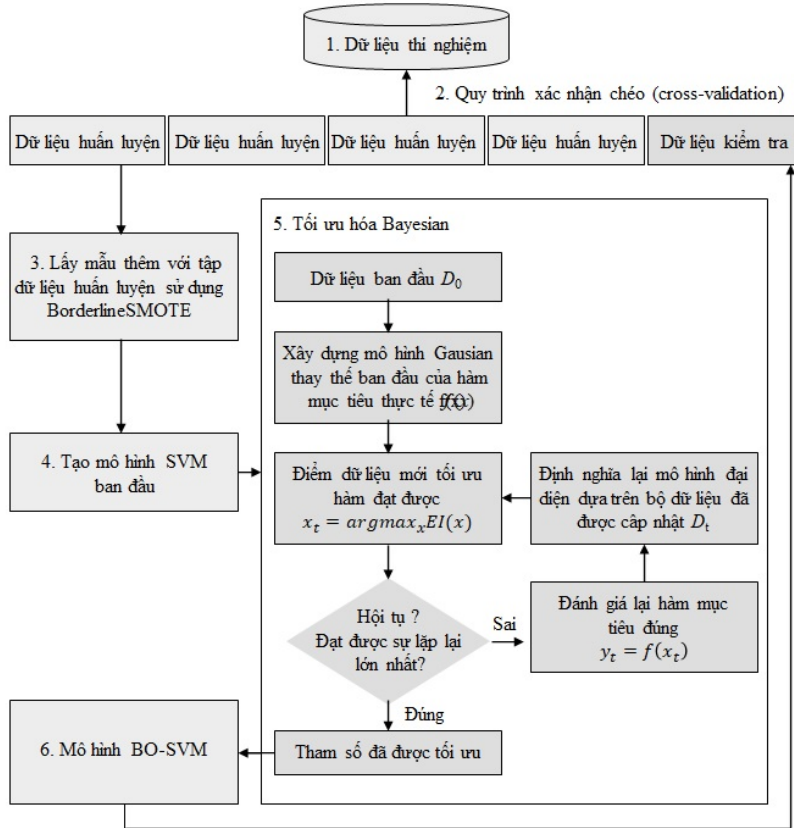
+ Bước 1: Chuẩn bị bộ dữ liệu thí nghiệm dùng để huấn luyện mô hình. Trong bước này, bộ dữ liệu đã được thu thập và xử lý trước theo phương pháp được trình bày trong phần “Tập dữ liệu thí nghiệm”.

+ Bước 2: Tách bộ dữ liệu bằng phương pháp xác nhận chéo k lần (k -fold cross-validation). Dữ liệu được chia thành năm phần phân tầng bao gồm một phần thử nghiệm và bốn phần để đào tạo mô hình. Trong quá trình xác nhận chéo (cross-validation) này, mô hình BO-SVM đã được đào tạo và thử nghiệm độc lập năm lần. Phần thử nghiệm lần lượt được thay thế bằng một phần khác sau mỗi lần lặp lại. Kết quả sẽ là trung bình của kết quả thử nghiệm hơn năm lần. Với bộ dữ liệu tương đối mất cân bằng trong nghiên cứu này, quá trình xác nhận chéo (cross-validation) đã giúp giảm sự dự đoán thiên lệch về một lớp nào đó cũng như giảm được vấn đề quá khớp (overfitting problems).

+ Bước 3: Lấy thêm mẫu với các lớp thiểu số. Trong bước này, số lượng mẫu trong mỗi lớp được cân bằng bằng phương pháp BorderlineSMOTE. Các điểm dữ liệu tổng hợp mới được tạo ra dựa trên mối quan hệ giữa các điểm hiện có.

+ Bước 4: Thiết lập thuật toán SVM ban đầu như là mô hình cơ sở.

+ Bước 5: Tối ưu hóa Bayesian. Bước này bao gồm quy trình tối ưu hóa cho ba tham số C , loại hàm “kernel” (kernel function) và tham số “kernel” γ sử dụng thuật toán BO. Khoảng khảo sát cho ba thông số này lần lượt là [1, 1000], [‘linear’, ‘poly’, ‘rbf’] và [1e-4, 10]. Những khoảng khảo sát này được chọn sau khi thực hiện một số quy trình tối ưu hóa đầu tiên để tìm ra phạm vi có thể có và phù hợp của các tham số. Cần lưu ý rằng do sử dụng quy trình xác nhận chéo (cross-validation), sẽ có năm bộ tham số tối ưu có thể đạt được. Tuy nhiên, chỉ có giá trị tốt nhất được chọn để xây dựng mô hình cuối cùng. Do sự mất cân bằng của bộ dữ liệu, tiêu chí đánh giá được chọn là tối đa hóa điểm số $F1$ ($F1$ -score) thay vì tối thiểu hóa mất mát (loss) vì nó thường gây ra sự thiên lệch đối với lớp đa số. Quá trình tối ưu BO lặp lại 5 lần với mỗi mô hình để chọn được các thông số tối ưu.



Hình 3. Thuật toán mô hình BO-SVM đề xuất để dự mức mức độ phá hoại cục bộ của tấm FRC do tải trọng va đập của đầu đạn gây ra

+ Bước 6: Xây dựng mô hình BO-SVM bằng cách sử dụng tham số tối ưu thu được. Sau đó, mô hình cuối cùng được thử nghiệm trên phần thử nghiệm được giữ lại.

Sau đó, quy trình được lặp lại từ Bước 2, nơi một bộ dữ liệu huấn luyện khác được tạo ra bởi quy trình xác nhận chéo (cross-validation). Toàn bộ quy trình của mô hình được đề xuất được thực hiện bằng ngôn ngữ lập trình Python.

4. Kết quả và thảo luận

Trong phần này, kết quả của các mô hình dự đoán mức độ phá hoại cục bộ của tấm bê tông cốt sợi chịu tác động va đập của đầu đạn sẽ được phân tích và đánh giá. Hiệu quả dự đoán của mô hình đề xuất được đánh giá thông qua các chỉ số: đường cong đặc trưng hoạt động của bộ thu nhận (Receiver Operating Characteristic - ROC), diện tích dưới đường cong ROC (Area Under ROC curve - AUC), điểm số $F1$ ($F1$ -score và cuối cùng là mức độ dự đoán chính xác lớp (class accuracy). Thời gian tính toán của các mô hình tương đối ngắn và không có sự chênh lệch giữa các mô hình do tập dữ liệu không quá lớn, vì vậy tiêu chí thời gian không được đề cập trong nghiên cứu này. Trong khi AUC được xác định bằng diện tích nằm giữa đường cong ROC và trục tỷ lệ khẳng định sai (trục hoành trong đồ thị đường cong ROC) trong phạm vi từ 0 đến 1 thì điểm số $F1$ được xác định như sau:

$$F1 = 2 \frac{Precision * Recall}{Precision + Recall} \quad (7)$$

trong đó: $Precision = \frac{TP}{TP + FP}$ và $Recall = \frac{TP}{TP + FN}$ với TP là tỷ lệ dự đoán đúng lớp khẳng định đúng (True Positive), TN là tỷ lệ dự đoán đúng lớp khẳng định sai (True negative), FP là tỷ lệ dự sai lớp khẳng định đúng (False Positive), và FN là tỷ lệ dự đoán sai lớp khẳng định sai (False Negative).

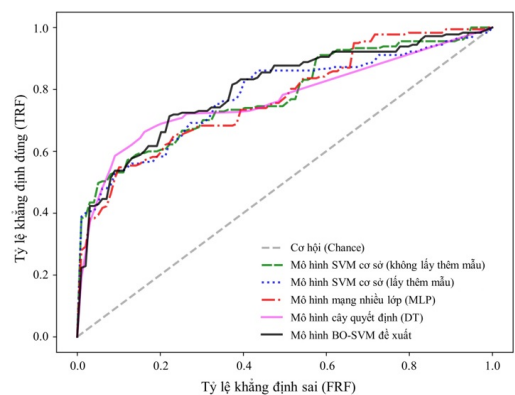
Các kết quả này từ mô hình đề xuất BO-SVM cũng được so sánh với các kết quả tương ứng thu được từ một số thuật toán máy học phổ biến khác như cây quyết định (Decision tree - DT), mạng nhiều lớp (Multi-layer Perceptron - MLP), mô hình SVM không được tối ưu hóa hay còn gọi là mô hình SVM cơ sở trong cả hai trường hợp có và không sử dụng kỹ thuật lấy thêm mẫu BorderlineSMOTE. Tất cả các siêu tham số được chọn cho các mô hình trên đều được lựa chọn thông qua quá trình thử dần (trial-and-error) một cách cẩn thận để tránh vấn đề quá khớp. Các thông số tối ưu của các mô hình (các siêu tham số thu được đã cho kết quả tốt nhất trong mỗi mô hình) được thể hiện trong Bảng 3.

Bảng 3. Siêu tham số tối ưu sử dụng cho các mô hình

Mô hình	Thông số sử dụng
Mô hình SVM cơ sở (không lấy thêm mẫu)	$C = 585$, loại hàm $kernel = \text{"linear kernel"}$, tham số "kernel" $\gamma = 3,96$
Mô hình SVM cơ sở (lấy thêm mẫu)	$C = 340$, loại hàm $kernel = \text{"radial basis function kernel"}$ (hoặc rbf kernel), tham số "kernel" $\gamma = 0,53$
Mô hình mạng nhiều lớp (MLP)	Số lớp ẩn " $number_of_hidden_layer$ " = 1, số lượng mạng " $number_of_neurons$ " = 20, " $l2_regularization$ " = 0,001, " $batch_size$ " = 16, " $learning_rate$ " = 0,084, " $activation_function$ " = 'tanh'
Mô hình cây quyết định (DT)	" max_depth " = 15, " $criterion$ " = 'entropy', " $min_samples_split$ " = 7,0, " $min_samples_leaf$ " = 3,0
Mô hình BO-SVM đề xuất	$C = 352$, loại hàm $kernel = \text{"radial basis function kernel"}$, tham số $\gamma = 1,61$

Hình 4 thể hiện đường cong ROC (The receiver operating characteristic curve) dựa trên tỷ lệ khẳng định đúng (True positive rate - TPR) và tỷ lệ khẳng định sai (False positive rate - FPR) thu được từ các mô hình. Đường cong ROC là một tiêu chí đánh giá nổi tiếng cho vấn đề đa lớp, nó cho biết khả năng phân loại của các mô hình giữa các lớp. Như có thể thấy từ Hình 4, đường cong ROC của mô hình BO-SVM đề xuất bao phủ hầu hết các mô hình khác. Đánh giá định lượng của các đường cong này được thể hiện thông qua diện tích phần dưới đường cong, cụ thể là AUC. AUC đại diện cho phép đo tổng hợp về độ ổn định và hiệu suất của mô hình trên tất cả các ngưỡng phân loại.

AUC càng cao chứng tỏ mô hình có khả năng dự đoán càng tốt. Các giá trị thu được của AUC của mỗi mô hình đã được thể hiện trong Bảng 4. Có thể thấy, mô hình BO-SVM được đề xuất có giá trị AUC cao nhất bằng 80,0%. Trong khi mô hình SVM cơ sở lấy thêm mẫu và không lấy thêm mẫu có giá



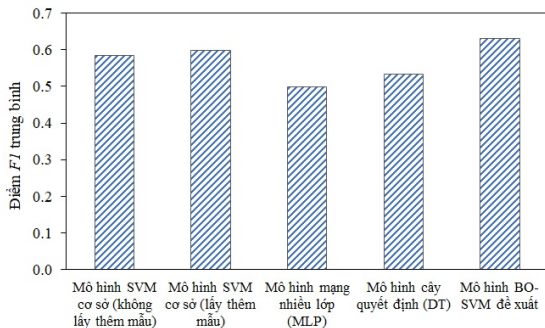
Hình 4. Đường cong ROC của các mô hình

trị AUC thấp hơn với lần lượt là 78,1% và 77,9%. AUC của mô hình BO-SVM cũng cao hơn các mô hình chuẩn khác bao gồm mô hình mạng nhiều lớp MLP (AUC = 77,2%) và mô hình cây quyết định DT (AUC = 77,0%). Những kết quả này cho thấy khả năng dự đoán cao hơn của mô hình BO-SVM so với các mô hình được chọn khác về khả năng phân loại mức độ phá hoại của tấm FRC do tác động va đập của đầu đạn gây ra.

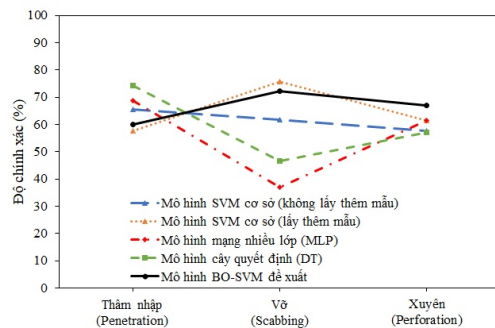
Bảng 4. Kết quả AUC của các mô hình

Mô hình	AUC (%)
Mô hình SVM cơ sở (không lấy thêm mẫu)	77,9 ± 22,2
Mô hình SVM cơ sở (lấy thêm mẫu)	78,1 ± 26,7
Mô hình mạng nhiều lớp (MLP)	77,2 ± 19,1
Mô hình cây quyết định (DT)	77,0 ± 18,9
Mô hình BO-SVM đề xuất	80,0 ± 19,7

Bên cạnh đó, kết quả điểm số *F1* trung bình trên 5 phần dữ liệu được thể hiện trong Hình 5. Có thể quan sát từ biểu đồ rằng, điểm *F1* của mô hình BO-SVM cao hơn các mô hình khác. Đặc biệt, mô hình BO-SVM đạt được điểm *F1* = 0,632, trong khi mô hình SVM cơ sở lấy thêm mẫu và không lấy thêm mẫu, mô hình mạng nhiều lớp MLP và mô hình cây ra quyết định DT đạt điểm *F1* lần lượt tương ứng là 0,584; 0,598; 0,499 và 0,534. Kết quả này đã thể hiện khả năng dự đoán tốt hơn của mô hình BO-SVM trên bộ dữ liệu thu được từ thí nghiệm tấm FRC chịu tác động va đập của đầu đạn.



Hình 5. Giá trị trung bình điểm số *F1* của các mô hình phân tích



Hình 6. Khả năng dự đoán mức độ phá hoại cục bộ tấm FRC của các mô hình

Bảng 5 và Hình 6 thể hiện kết quả độ chính xác của các mô hình trong dự đoán mức độ phá hoại cục bộ của các trường hợp phá hoại cục bộ của tấm FRC dưới tác động va đập của đầu đạn. Có thể thấy từ Bảng 4 và Hình 6, chỉ có mô hình SVM cơ sở không sử dụng kỹ thuật lấy thêm mẫu cho kết quả dự đoán các trường hợp phá hoại cục bộ đồng đều hơn mô hình BO-SVM đề xuất, tuy nhiên kết quả dự đoán lại kém chính xác hơn. Các mô hình còn lại khác đều cho kết quả dự đoán các trường hợp phá hoại cục bộ kém đồng đều cũng như kém chính xác hơn mô hình BO-SVM đề xuất. Cụ thể, độ chính xác trung bình trong dự đoán mức độ phá hoại với các trường hợp từ mô hình BO-SVM là 66,5%, cao hơn kết quả tương ứng thu được từ mô hình SVM cơ bản lấy thêm mẫu và không lấy thêm mẫu với kết quả lần lượt là 65,0% và 61,8%, cũng như cao hơn kết quả tương ứng thu được từ mô hình mạng nhiều lớp MLP và mô hình cây ra quyết định DT với kết quả lần lượt là 55,8% và 59,4%, như được thể hiện trong Bảng 5.

Bảng 5. Khả năng dự đoán mức độ phá hoại cục bộ của tấm FRC của các mô hình

Mô hình	Lớp 1 (Penetration) (%)	Lớp 2 (Scabbing) (%)	Lớp 3 (Perforation) (%)	Trung bình lớp (%)
Mô hình SVM (không lấy thêm mẫu)	65,6	61,9	57,8	61,8
Mô hình SVM (lấy thêm mẫu)	57,8	75,7	61,5	65,0
Mô hình mạng nhiều lớp (MLP)	68,9	37,1	61,5	55,8
Mô hình cây quyết định (DT)	74,4	46,7	57,1	59,4
Mô hình BO-SVM đề xuất	60,1	72,4	66,9	66,5

Ngoài ra, cũng cần lưu ý rằng mô hình BO-SVM cho kết quả độ chính xác trong dự đoán trường hợp phá hoại cục bộ vỡ (Scabbing) là lớn nhất (mức độ chính xác 72,4%) mặc dù số dữ liệu thí nghiệm trong trường hợp này là nhỏ nhất (33 mẫu). Việc lớp vỡ (Scabbing) được dự đoán với độ chính xác cao nhất là do kỹ thuật lấy thêm mẫu đã tạo ra nhiều mẫu dữ liệu tổng hợp chỉ trong phạm vi một số lượng nhỏ mẫu của lớp này. Trong quá trình kiểm tra, khi một mẫu mới được đưa vào thì mô hình sẽ xem xét mẫu đó với các đường biên đã được tạo ra từ trước, từ đó cho ra kết quả tốt hơn các mô hình phân loại truyền thống như phân loại tuyến tính (Linear classification) hay phân loại KNN (*k*-nearest neighbors classification) ... đặc biệt với các dữ liệu có tính phi tuyến và chồng chéo lên nhau như trong nghiên cứu này. Kết quả trên cũng cho thấy sự hiệu quả và cần thiết của việc sử dụng thêm kỹ thuật lấy thêm mẫu BorderlineSMOTE trong xây dựng mô hình BO-SVM.

Mặc dù độ chính xác trong dự đoán dạng phá hoại thâm nhập (Penetration) của mô hình BO-SVM đề xuất thấp hơn một số mô hình khác nhưng độ chính xác trung bình của mô hình đề xuất vẫn cao hơn. Hơn nữa, về cơ bản độ chính xác trung bình của ba mức độ phá hoại trong mô hình BO-SVM đề xuất cân bằng và cao hơn so với các mô hình dự đoán khác. Điều này rất quan trọng khi dự đoán mức độ phá hoại cục bộ của tấm FRC dưới tác động va đập của đầu đạn, là bài toán mà các thông số đầu vào thiết kế gần như không thể tính toán hoặc dự đoán chính xác được.

5. Kết luận

Trong bài báo này, một thuật toán máy học mạnh Support vector machine (SVM) được sử dụng để xây dựng mô hình phân loại cho việc dự đoán mức độ phá hoại cục bộ của tấm bê tông cốt sợi FRC dưới tác động va đập của đầu đạn. Những siêu tham số được tối ưu bằng cách sử dụng phương pháp tối ưu Bayesian (Bayesian optimization-BO). Mô hình có chứa các tham số đã được tối ưu để dự đoán mức độ phá hoại cục bộ của tấm FRC gọi là BO-SVM. Từ kết quả nghiên cứu, một số kết luận được rút ra như sau:

- Mô hình BO-SVM đề xuất được huấn luyện trên bộ dữ liệu gồm 176 thí nghiệm tấm FRC chịu tác động va đập của đầu đạn cho giá trị AUC bằng 80,0%, là kết quả tốt nhất so với các mô hình chuẩn cơ bản gồm SVM cơ sở lấy thêm mẫu và không lấy thêm mẫu, mạng nhiều lớp MLP và cây quyết định DT.

- Phương pháp tối ưu BO góp phần quan trọng và hiệu quả trong việc tìm ra các giá trị siêu tham số (hyperparameters) tốt nhất cho mô hình SVM đạt kết quả dự đoán cao hơn về các thông số điểm $F1$ (0,632) cũng như khả năng dự đoán chính xác (66,5%) so với các mô hình cơ bản khác.

- Kỹ thuật lấy thêm mẫu BorderlineSMOTE là cần thiết và hiệu quả trong việc xây dựng mô hình dự đoán mức độ phá hoại của tấm FRC nhằm đưa ra kết quả dự đoán cân bằng, đặc biệt trong trường hợp các dữ liệu có tính phi tuyến và chồng chéo lên nhau như trong nghiên cứu này.

- Mô hình đề xuất BO-SVM có thể được sử dụng như một công cụ trong thực hành thiết kế để dự đoán mức độ phá hoại cục bộ của tấm FRC dưới tác động va đập của đầu đạn, đặc biệt trong giai đoạn thiết kế ban đầu.

Kết quả dự đoán của mô hình BO-SVM đề xuất tốt hơn các mô hình cơ bản khác, tuy nhiên mức độ dự đoán chính xác của mô hình là chưa cao thực sự (trung bình cho 3 lớp là 66,5%). Một lý do khác quan là do các dữ liệu được tập hợp từ nhiều nguồn khác nhau với rất nhiều thông số đầu vào khác nhau, không hội tụ thành từng nhóm, do đó kết quả dự đoán như trên có thể coi là chấp nhận được với bộ số liệu này. Mô hình dự đoán có thể phát triển để nâng cao hiệu suất nếu bộ dữ liệu được thu thập phong phú hơn cũng như sử dụng thêm các kỹ thuật lấy mẫu thêm mạnh hơn, các nội dung này có thể được xem xét trong các nghiên cứu tiếp theo.

Lời cảm ơn

Nghiên cứu này được tài trợ bởi Bộ Giáo dục và Đào tạo thông qua Đề tài Khoa học và Công nghệ cấp Bộ, mã số: B2022-XDA-05. Một số kết quả được thực hiện thông qua Đề tài nghiên cứu Khoa học sinh viên năm học 2021-2022, Trường Đại học Xây dựng Hà Nội, mã số: XD-2022-03.

Tài liệu tham khảo

- [1] Kennedy, R. P. (1976). [A review of procedures for the analysis and design of concrete structures to resist missile impact effects](#). *Nuclear Engineering and Design*, 37(2):183–203.
- [2] Hashimoto, J., Takiguchi, K., Nishimura, K., Matsuzawa, K., Tsutsui, M., Ohashi, Y., Kojima, I., Torita, H. (2005). Experimental study on behavior of RC panels covered with steel plates subjected to missile impact. *18th International Conference on Structure Mechanics in Reactor Technology (SMiRT 18)*, IASMiRT, 2604–2615.
- [3] Li, Q. M., Reid, S. R., Wen, H. M., Telford, A. R. (2005). [Local impact effects of hard missiles on concrete targets](#). *International Journal of Impact Engineering*, 32(1-4):224–284.
- [4] Kim, J. J., Park, G.-J., Kim, D. J., Moon, J. H., Lee, J. H. (2014). [High-rate tensile behavior of steel fiber-reinforced concrete for nuclear power plants](#). *Nuclear Engineering and Design*, 266:43–54.
- [5] Park, J. K., Kim, S.-W., Kim, D. J. (2017). [Matrix-strength-dependent strain-rate sensitivity of strain-hardening fiber-reinforced cementitious composites under tensile impact](#). *Composite Structures*, 162: 313–324.
- [6] Dancygier, A. N., Yankelevsky, D. Z. (1996). [High strength concrete response to hard projectile impact](#). *International Journal of Impact Engineering*, 18(6):583–599.
- [7] Luo, X., Sun, W., Chan, S. Y. N. (2001). [Steel fiber reinforced high-performance concrete: a study on the mechanical properties and resistance against impact](#). *Materials and Structures*, 34(3):144–149.
- [8] Zhang, M. H., Shim, V. P. W., Lu, G., Chew, C. W. (2005). [Resistance of high-strength concrete to projectile impact](#). *International Journal of Impact Engineering*, 31(7):825–841.
- [9] Jamnam, S., Maho, B., Techaphatthanakon, A., Sonoda, Y., Yoo, D.-Y., Sukontasukkul, P. (2020). [Steel fiber reinforced concrete panels subjected to impact projectiles with different caliber sizes and muzzle energies](#). *Case Studies in Construction Materials*, 13:e00360.
- [10] Soe, K. T., Zhang, Y. X., Zhang, L. C. (2013). [Impact resistance of hybrid-fiber engineered cementitious composite panels](#). *Composite Structures*, 104:320–330.
- [11] Li, Q. M., Chen, X. W. (2003). [Dimensionless formulae for penetration depth of concrete target impacted by a non-deformable projectile](#). *International Journal of Impact Engineering*, 28(1):93–116.
- [12] Li, Q. M., Tong, D. J. (2003). [Perforation thickness and ballistic limit of concrete target subjected to rigid projectile impact](#). *Journal of Engineering Mechanics*, 129(9):1083–1091.
- [13] Kosteski, L. E., Riera, J. D., Iturrioz, I., Singh, R. K., Kant, T. (2015). [Assessment of empirical formulas for prediction of the effects of projectile impact on concrete structures](#). *Fatigue & Fracture of Engineering Materials & Structures*, 38(8):948–959.

- [14] Almusallam, T. H., Siddiqui, N. A., Iqbal, R. A., Abbas, H. (2013). [Response of hybrid-fiber reinforced concrete slabs to hard projectile impact](#). *International Journal of Impact Engineering*, 58:17–30.
- [15] Almusallam, T. H., Abadel, A. A., Al-Salloum, Y. A., Siddiqui, N. A., Abbas, H. (2015). [Effectiveness of hybrid-fibers in improving the impact resistance of RC slabs](#). *International Journal of Impact Engineering*, 81:61–73.
- [16] Hùng, H. M., Hùng, T. V., Thuật, Đ. V., Việt, V. Q. (2020). [Phương pháp xử lý hiện tượng mô hình quá khớp trong xây dựng mô hình học sâu để ước lượng khả năng chịu tải của giàn phi tuyến](#). *Tạp chí Khoa học Công nghệ Xây dựng (KHCN XD) - ĐHXDHN*, 14(1V):12–20.
- [17] Hoàn, P. T. (2021). [Ước lượng khả năng chịu nén đúng tâm của cột ống thép nhồi bê tông bằng thuật toán máy học](#). *Tạp chí Khoa học Công nghệ Xây dựng (KHCN XD) - ĐHXDHN*, 15(3V):69–78.
- [18] Thai, D.-K., Tu, T. M., Bui, T. Q., Bui, T.-T. (2019). [Gradient tree boosting machine learning on predicting the failure modes of the RC panels under impact loads](#). *Engineering with Computers*, 37(1):597–608.
- [19] Thai, D.-K., Nguyen, D.-L., Pham, T.-H., Doan, Q. H. (2021). [Prediction of residual strength of FRC columns under blast loading using the FEM method and regression approach](#). *Construction and Building Materials*, 276:122253.
- [20] Doan, Q. H., Le, T., Thai, D.-K. (2021). [Optimization strategies of neural networks for impact damage classification of RC panels in a small dataset](#). *Applied Soft Computing*, 102:107100.
- [21] Pham, T.-H., Nguyen, N.-V. (2021). [Mechanical properties of constituent phases in structural steels and heat-affected zones investigated by statistical nanoindentation analysis](#). *Construction and Building Materials*, 268:121211.
- [22] Truong, V.-H., Ly, Q. V., Le, V.-C., Vu, T.-B., Le, T.-T.-T., Tran, T.-T., Goethals, P. (2021). [Machine learning-based method for forecasting water levels in irrigation and drainage systems](#). *Environmental Technology & Innovation*, 23:101762.
- [23] Vu, Q.-V., Truong, V.-H., Thai, H.-T. (2021). [Machine learning-based prediction of CFST columns using gradient tree boosting algorithm](#). *Composite Structures*, 259:113505.
- [24] Kim, S.-E., Vu, Q.-V., Papazafeiropoulos, G., Kong, Z., Truong, V.-H. (2020). [Comparison of machine learning algorithms for regression and classification of ultimate load-carrying capacity of steel frames](#). *Steel and Composite Structures*, 37(2):193–209.
- [25] ACI 349-01 (2001). *Code requirements of nuclear safety related concrete structures*. American Concrete Institute.
- [26] Dancygier, A. N., Yankelevsky, D. Z., Jaegermann, C. (2007). [Response of high performance concrete plates to impact of non-deforming projectiles](#). *International Journal of Impact Engineering*, 34(11): 1768–1779.
- [27] Riedel, W., Nöldgen, M., Straßburger, E., Thoma, K., Fehling, E. (2010). [Local damage to ultra high performance concrete structures caused by an impact of aircraft engine missiles](#). *Nuclear Engineering and Design*, 240(10):2633–2642.
- [28] Hossin, M., Sulaiman, M. N. (2015). [A review on evaluation metrics for data classification evaluations](#). *International Journal of Data Mining & Knowledge Management Process*, 5(2):01–11.
- [29] Pelikan, M., Goldberg, D. E., Cantú-Paz, E. (1999). BOA: The Bayesian optimization algorithm. *Proceedings of the genetic and evolutionary computation conference GECCO-99*, Citeseer, 1:525–532.
- [30] Truong, V.-H., Pham, H.-A. (2021). [Support vector machine for regression of ultimate strength of trusses: A comparative study](#). *Engineering Journal*, 25(7):157–166.
- [31] Han, H., Wang, W.-Y., Mao, B.-H. (2005). [Borderline-SMOTE: A new over-sampling method in imbalanced data sets learning](#). *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, 878–887.
- [32] Chawla, N. V., Bowyer, K. W., Hall, L. O., Kegelmeyer, W. P. (2002). [SMOTE: Synthetic Minority Over-sampling Technique](#). *Journal of Artificial Intelligence Research*, 16:321–357.
- [33] Santos, M. S., Soares, J. P., Abreu, P. H., Araujo, H., Santos, J. (2018). [Cross-validation for imbalanced datasets: avoiding overoptimistic and overfitting approaches](#). *IEEE Computational Intelligence Magazine*, 13(4):59–76.