

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

Hà Mạnh Hùng^a, Trương Việt Hùng^{b,*}, Đinh Văn Thuật^a, Vũ Quang Việt^c

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

^b*Khoa Công trình, Trường Đại học Thủy Lợi, 175 đường Tây Sơn, quận Đống Đa, Hà Nội, Việt Nam*

^c*Khoa Công trình, Trường Đại học Hàng Hải Việt Nam, 484 đường Lạch Tray, quận Lê Chân, Hải Phòng, Việt Nam*

Nhận ngày 13/10/2019, Sửa xong 02/11/2019, Chấp nhận đăng 22/11/2019

Tóm tắt

Hiện tượng mô hình quá khớp (overfitting) là một trong những nguyên nhân quan trọng làm giảm hiệu quả của mô hình học sâu, đặc biệt là trong các bài toán có tính phi tuyến cao như bài toán ước lượng khả năng chịu tải của giàn làm bằng kim loại. Bài báo này sẽ trình bày, phân tích và so sánh hiệu quả của một số kỹ thuật thường được áp dụng hiện nay cho việc xử lý hiện tượng mô hình quá khớp bao gồm kỹ thuật dừng sớm (Early Stopping), nhớ mô hình (Model Checkpoint) và kết hợp hai kỹ thuật trên. Một giàn phẳng gồm 39 thanh được sử dụng để minh họa cho nghiên cứu. Tập dữ liệu cho mô hình học sâu được tạo ra từ phân tích phi tuyến giàn có thông số đầu vào là diện tích mặt cắt ngang của các thanh giàn và thông số đầu ra là hệ số khả năng chịu tải (LF). Kết quả nghiên cứu cho thấy phương pháp kết hợp cả hai kỹ thuật dừng sớm và nhớ mô hình đem lại hiệu quả cao nhất về cả góc độ thời gian huấn luyện và độ chính xác của mô hình.

Từ khóa: học sâu; trí tuệ nhân tạo; phân tích trực tiếp; kết cấu giàn; máy học.

METHODS TO SOLVE OVERFITTING OF DEEP LEARNING MODEL FOR ESTIMATING LOAD-CARRYING CAPACITY OF NONLINEAR TRUSSES

Abstract

Overfitting is one of the most important problems to reduce the performance of a deep learning model, especially in highly nonlinear problems such as estimation of load-carrying capacity of nonlinear inelastic steel trusses. This paper presents common methods to solve overfitting such as Early Stopping, Model Checkpoint, and the combination of both methods. A planar steel truss with 39 elements is used to illustrate the study. Data for the deep learning model is generated by using an advanced analysis, where the inputs are cross-sectional areas of the truss elements and output is the ultimate load factor of the structure. The results show that the combination of Early Stopping and Model Checkpoint provide the best performance regarding both computational efforts and accuracy.

Keywords: deep learning; artificial intelligent; advanced analysis; truss; machine learning.

[https://doi.org/10.31814/stce.nuce2020-14\(1V\)-02](https://doi.org/10.31814/stce.nuce2020-14(1V)-02) © 2020 Trường Đại học Xây dựng (NUCE)

1. Đặt vấn đề

Kết cấu giàn được sử dụng phổ biến hiện nay do những ưu điểm nổi trội của loại kết cấu này như vượt nhịp lớn, phát huy tối đa sự làm việc của vật liệu, hình thức đẹp, nhẹ, linh hoạt và phong phú.

*Tác giả chính. Địa chỉ e-mail: truongviethung@tlu.edu.vn (Hùng, T. V.)

Tương tự như những loại kết cấu thép khác, tính chất phi tuyến hình học và phi tuyến vật liệu cần được xét đến khi phân tích hệ giàn thép nhằm mô tả chính xác ứng xử của công trình. Để xét đến tính phi tuyến của kết cấu, các phương pháp thiết kế truyền thống được xây dựng dựa trên hai bước cơ bản là: (i) xác định nội lực của các cấu kiện của hệ kết cấu sử dụng phân tích tuyến tính đàn hồi và (ii) phân tích, đánh giá và thiết kế cho từng cấu kiện riêng lẻ của công trình có xét đến các ảnh hưởng phi tuyến bằng các công thức thiết kế được cung cấp trong các tiêu chuẩn hiện hành như AISC LRFD [1], Eurocode [2], ... Rõ ràng, cách tiếp cận này không mô tả được một cách trực tiếp các ứng xử phi tuyến của cả hệ kết cấu, đồng thời việc thiết kế riêng lẻ cho từng cấu kiện không đảm bảo sự tương thích của chúng với cả hệ kết cấu. Vì thế, mặc dù các phương pháp thiết kế này đơn giản và cho kết quả thiết kế chấp nhận được, chúng đang dần được thay thế bằng các phương pháp phân tích trực tiếp cho phép xét đến cả phi tuyến hình học và phi tuyến vật liệu của công trình [3–9].

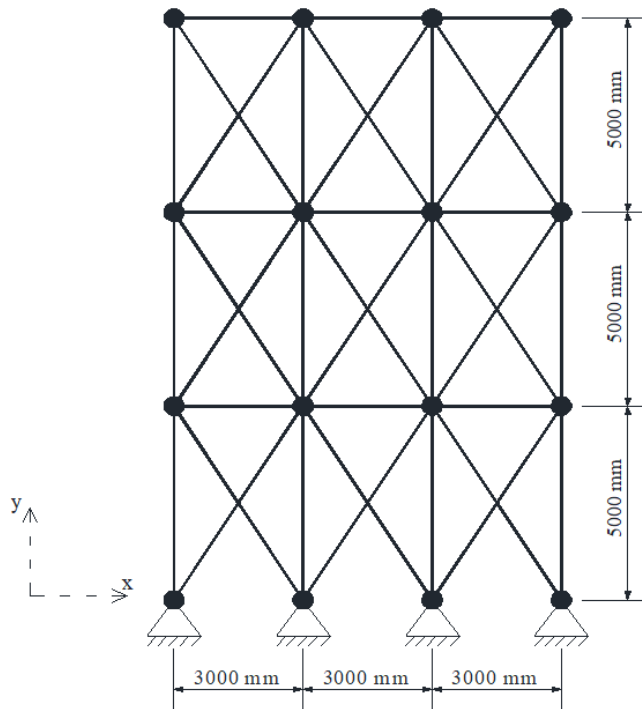
Một nhược điểm quan trọng của phương pháp phân tích trực tiếp đó là thời gian tính khá lâu, đặc biệt đối với các bài toán đòi hỏi số lần phân tích kết cấu là rất lớn như bài toán tối ưu, tính độ tin cậy của công trình, ... Để khắc phục vấn đề này, một hướng tiếp cận khá hay và thu hút được sự quan tâm, nghiên cứu của nhiều học giả là sử dụng thuật toán máy học (Machine Learning - ML). Một số nghiên cứu nổi bật về việc ứng dụng các thuật toán ML có thể kể đến như: Zhang và cs. [10], Breiman [11], Friedman [12], Safavian và Landgrede [13], Worden và Lane [14], ... Thuật toán học sâu (Deep Learning - DL) sử dụng mô hình mạng nơ-ron nhân tạo (Artificial Neural Network - ANN) là một trong những thuật toán mới nhất của ML. Thuật toán DL đã được áp dụng thành công trong rất nhiều lĩnh vực về thiết kế công trình như: phát hiện hư hỏng [15], quan trắc công trình [16], ... Các mô hình DL dựa trên mạng ANN được tổ chức bằng một hệ thống các lớp đại diện khác nhau (layer) được liên kết đơn giản nhưng phi tuyến với nhiều nơ-ron trên mỗi lớp. Các lớp sau có độ phức tạp cao hơn lớp trước. Thông qua hệ thống các lớp đại diện, các thông tin quan trọng của dữ liệu đầu vào sẽ được khuếch đại trong khi các thông tin kém quan trọng hoặc thông tin nhiễu sẽ được giảm dần và loại trừ khỏi hệ thống. Với một hệ thống tổ chức phức tạp nhưng linh hoạt như vậy, mô hình DL có khả năng xử lý các dữ liệu có độ phức tạp cao và nhiều chiều. Mặc dù việc xây dựng mô hình DL là khá đơn giản và không đòi hỏi kỹ thuật quá cao của người sử dụng, việc khảo sát đánh giá các tác động của các tham số của mô hình DL trong mỗi bài toán cụ thể là rất quan trọng. Đối với bài toán ước lượng khả năng chịu tải của kết cấu giàn phi tuyến, Hùng và cs. [17] gần đây đã tiến hành xây dựng mô hình DL và phân tích đánh giá ảnh hưởng của các tham số của mô hình. Nhiều kết quả đã được đưa ra, tuy nhiên các tác giả chưa xem xét đến vấn đề mô hình quá khớp (overfitting) trong nghiên cứu của mình. Điều này khiến cho mô hình được xây dựng chưa hoàn thiện và có nhiều hạn chế khi áp dụng vào thực tế.

Mô hình quá khớp là việc mô hình thu được sau khi huấn luyện có hiện tượng đạt hiệu quả quá tốt với tập dữ liệu huấn luyện nhưng hiệu quả lại không cao đối với một tập dữ liệu mới. Nguyên nhân của hiện tượng này có thể giải thích như sau. Các mô hình học sâu rất mạnh mẽ trong việc ước lượng các đặc tính của dữ liệu. Khi số vòng huấn luyện (epochs) càng lớn thì mô hình ước lượng các đặc tính của dữ liệu có độ chính xác càng cao. Tuy nhiên, lúc này sẽ xảy ra hiện tượng là mô hình ước lượng cả những đặc tính tổng quát cũng như cá thể của tập dữ liệu huấn luyện. Nghĩa là, mô hình được xây dựng chưa tổng quát với toàn bộ tập dữ liệu mà mang quá nhiều đặc tính cá thể của tập huấn luyện. Phân tích trên đây cũng cho thấy rằng hiện tượng mô hình quá khớp thường liên quan trực tiếp đến số lượng vòng huấn luyện được sử dụng. Khi số lượng vòng huấn luyện quá lớn, mô hình sẽ quá khớp với tập huấn luyện nên giảm tính tổng quát của mô hình và dẫn đến hiệu quả mô hình giảm đi.

Bài báo này sẽ tập trung phân tích các phương pháp xử lý hiện tượng mô hình quá khớp cho bài toán xây dựng mô hình DL để ước lượng khả năng chịu tải của giàn phi tuyến. Một giàn phẳng gồm

39 thanh được sử dụng để minh họa cho nghiên cứu. Đầu tiên, phân tích trực tiếp được sử dụng để ước lượng ứng xử phi tuyến của kết cấu. Sau đó, tập dữ liệu sử dụng cho mô hình DL được xây dựng từ các kết quả thu được bằng việc sử dụng phân tích trực tiếp. Thông số đầu vào của mô hình là diện tích mặt cắt ngang của các phần tử giàn và thông số đầu ra là hệ số khả năng chịu tải (LF). LF là tỉ số của khả năng chịu tải lớn nhất của cả hệ kết cấu chia cho hiệu ứng của tải trọng. Các thông số cơ bản của mô hình được lựa chọn dựa trên khuyến nghị được đưa ra bởi Hùng và cs. [17]. Dựa trên mô hình DL trên, các kỹ thuật xử lý hiện tượng mô hình quá khớp được xem xét và so sánh.

2. Sơ đồ giàn phẳng 39 thanh và hiện tượng quá khớp



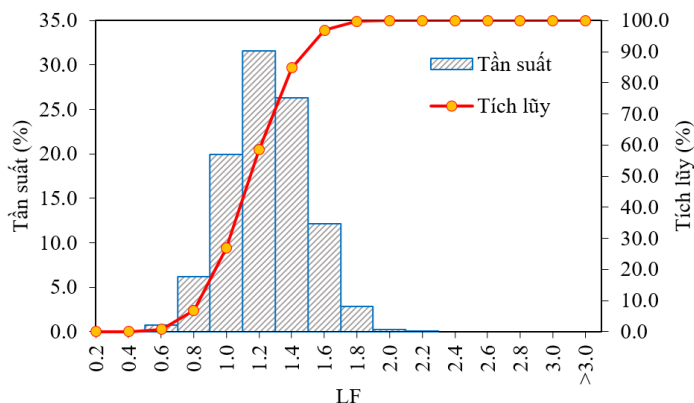
Hình 1. Giàn thép phẳng gồm 39 thanh

Trong bài báo này, một giàn phẳng có 39 thanh như Hình 1 được sử dụng với nhịp là 3 m và chiều cao là 5 m. Diện tích tiết diện của 39 thanh giàn được xem là 39 biến thiết kế và có thể nhận giá trị tùy ý trong khoảng từ 645,16 mm² đến 11290,3 mm². Vật liệu bằng nhôm có mô đun đàn hồi bằng 68950 MPa và cường độ chảy bằng 172,375 MPa. Tải trọng tác dụng theo phương ngang của giàn được qui đổi về tải trọng tập trung tại tất cả các nút và có giá trị bằng 136 kN. Tải trọng thẳng đứng cũng được qui đổi về tải trọng tập trung tại nút và có giá trị bằng 170 kN tác dụng theo chiều từ trên xuống. Tải trọng bản thân của giàn được bỏ qua. Biến đầu ra của mô hình là hệ số khả năng chịu tải của kết cấu (LF). Hệ số LF được tính dựa theo công thức sau:

$$LF = \frac{R}{S} \quad (1)$$

trong đó R và S tương ứng là khả năng chịu tải của công trình và tác dụng của tải trọng. LF được xác định bằng phần mềm phân tích phi tuyến kết cấu PAAP [18].

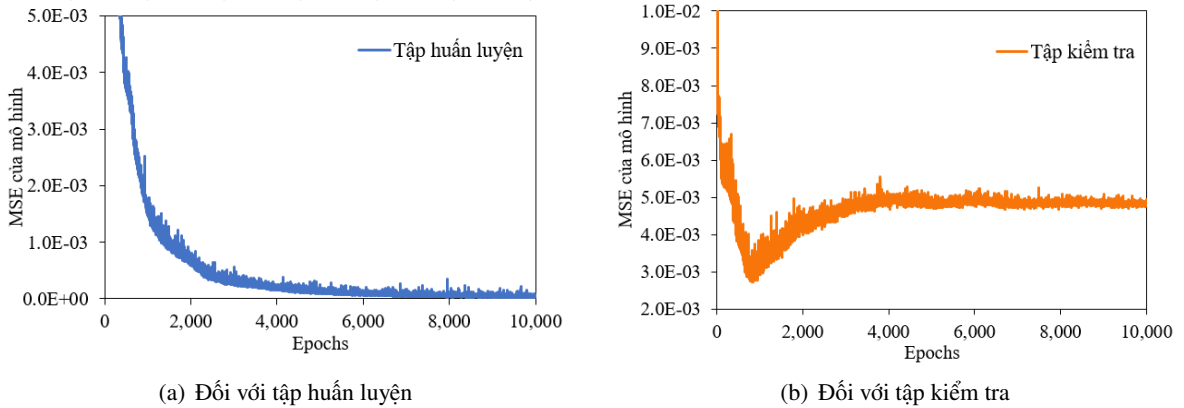
Để phục vụ cho nghiên cứu này, 50000 dữ liệu đã được tạo ra có biểu đồ phân bố xác suất như trong Hình 2. Hình 2 cho thấy rằng LF của 50000 dữ liệu phần lớn tập trung trong khoảng giá trị từ 0,6 đến 1,8 với trên 99%. Tỷ lệ dữ liệu có giá trị LF bé hơn 1 chiếm 26,9% nghĩa là xác suất công trình bị hư hỏng khi cho các dữ liệu tùy biến là 26,9%.



Hình 2. Biểu đồ xác suất của LF của giàn với 50000 dữ liệu

Mô hình DL dựa trên mạng ANN được xây dựng sử dụng ngôn ngữ lập trình Python và thư viện Keras và dựa trên kết quả nghiên cứu được trình bày trong tài liệu [17] với các thông số chính như sau. Mô hình lựa chọn là mô hình học có giám sát do dữ liệu đã có thông số đầu ra chính xác. 39 biến thiết kế trên được mô hình thành 39 biến đầu vào (input) trong mô hình DL. Nhằm tăng hiệu quả hội tụ của mô hình DL, các biến đầu vào được chuẩn hóa để giá trị nằm trong đoạn từ 0 đến 1 bằng cách chia cho giá trị 11290,3 mm². Lớp đầu ra gồm 1 nút tương ứng với giá trị LF. Số lượng lớp ẩn được chọn là 3 lớp với số nút trên các lớp là 64-128-64 đảm bảo nằm trong khoảng giá trị từ 1 đến 3 lần số biến đầu vào được khuyến nghị trong tài liệu [17]. Như vậy, cấu trúc của mô hình là 39-64-128-64-1. Hàm kích hoạt (activation function) được sử dụng là hàm LeakyReLU do ưu điểm hội tụ nhanh và kết quả khá tốt của nó. Tỷ lệ học (learning rate) được lấy bằng 0,01 và kích thước batch size là 128. Thuật toán tối ưu Adam được sử dụng cho quá trình huấn luyện do khả năng mạnh mẽ của thuật toán này để vượt qua các điểm tối ưu cục bộ để đạt đến điểm tối ưu nhất. Hàm tối ưu được lựa chọn cho quá trình huấn luyện là hàm ước lượng trung bình bình phương (mean square error - MSE). Chi tiết việc xây dựng mô hình ở trên có thể tham khảo trong tài liệu [17].

Hình 3 mô tả MSE của mô hình qua các lần lặp với số dữ liệu cho tập huấn luyện và tập kiểm tra cùng là 5000. Từ Hình 3(a) chúng ta có thể thấy khi số vòng lặp (Epochs) tăng, mô hình thu được cho sai số đối với tập huấn luyện càng giảm, tức là độ chính xác của mô hình cho việc ước lượng đầu ra của dữ liệu trong tập huấn luyện tăng lên. Tuy nhiên, Hình 3(b) lại cho thấy khi Epochs quá lớn (trong trường hợp này khi Epochs > 1.000) MSE của mô hình thu được đối với tập kiểm tra lại giảm dần. Điều này có nghĩa là độ chính xác của mô hình trong việc ước lượng các dữ liệu không được sử dụng trong tập huấn luyện lại bị giảm. Nguyên nhân của hiện tượng này là do khi Epochs khá lớn, mô hình thu được phản ánh rất tốt các đặc tính của tập huấn luyện bao gồm cả các đặc tính mang tính cá thể, cục bộ. Điều này dẫn đến khả năng dự đoán các đặc tính chung của toàn bộ dữ liệu của mô hình bị giảm, và kết quả là sai số của mô hình khi dự đoán các dữ liệu mới tăng lên. Vấn đề này được gọi là hiện tượng mô hình quá khớp (overfitting). Các phương pháp dùng để xử lý vấn đề này sẽ được trình bày trong các phần tiếp theo.



Hình 3. MSE của mô hình qua các lần lặp với dữ liệu tập huấn luyện và tập kiểm tra là 5000

3. Các kỹ thuật xử lý hiện tượng mô hình quá khớp

Để xử lý hiện tượng mô hình quá khớp có khá nhiều kỹ thuật có thể được áp dụng như là: kỹ thuật chính quy hóa (regularization), kỹ thuật tắt ngẫu nhiên một số nút trong mô hình (dropout), kỹ thuật kiểm tra chéo (cross-validation), ... Các kỹ thuật này thường đòi hỏi người dùng có một nền tảng kiến thức sâu về việc lập và xây dựng các mô hình DL. Xét dưới góc độ giúp cho người sử dụng có thể dễ hiểu và nhanh chóng áp dụng được trong công việc nghiên cứu của mình, trong phạm vi bài báo này chúng tôi sẽ tập trung giới thiệu hai thuật toán có tính hiệu quả rất cao nhưng lại khá dễ hiểu là kỹ thuật dừng sớm (Early Stopping - ES) và kỹ thuật nhớ mô hình (Model Checkpoint - MC).

3.1. Kỹ thuật dừng sớm (Early Stopping)

Dựa vào Hình 3 có thể thấy rằng, khi Epochs quá lớn sẽ dẫn đến độ chính xác của mô hình đối với tập kiểm tra bị giảm đi. Như vậy, một cách đơn giản nhất để xử lý vấn đề này là chúng ta cần tìm ra một thời điểm thích hợp để dừng việc huấn luyện lại để đảm bảo mô hình thu được cho sai số đối với tập kiểm tra là nhỏ nhất. Kỹ thuật này gọi là kỹ thuật dừng sớm (Early Stopping - ES). Để thực hiện kỹ thuật ES, dữ liệu cần chia thành 3 tập khác nhau là tập huấn luyện, tập kiểm định (validation set) và tập kiểm tra. Trong quá trình huấn luyện, độ chính xác của mô hình thu được sẽ liên tục được đánh giá dựa trên tập kiểm định nhằm tìm ra thời điểm dừng hợp lý. Tính hiệu quả của mô hình thu được sẽ được đánh giá bằng tập kiểm tra.

Kỹ thuật ES được thực hiện bằng cách sử dụng hàm EarlyStopping trong thư viện Keras. Dạng đơn giản nhất của hàm EarlyStopping như sau:

$$es1 = \text{EarlyStopping}(\text{monitor}='val_loss', \text{mode}='min') \quad (2)$$

trong đó *es1* là tên hàm do người dùng định nghĩa để sử dụng hàm EarlyStopping; đối số *monitor* dùng để chỉ ra chỉ số chúng ta muốn theo dõi trong quá trình huấn luyện để kết thúc sớm việc huấn luyện; *val_loss* sử dụng ở đây nghĩa là chỉ số chúng ta muốn theo dõi là MSE của tập kiểm định; đối số *mode* dùng để chỉ ra mục tiêu theo dõi và thường dùng với hai trường hợp là *min* và *max* tương ứng với giá trị nhỏ nhất và giá trị lớn nhất của MSE của tập kiểm định. Trong trường hợp nghiên cứu của bài báo, hàm tối ưu là MSE nên ta sử dụng chỉ số *min*. Khi sử dụng lệnh EarlyStopping này, chương trình huấn luyện sẽ ngay lập tức dừng lại nếu như trong vòng lặp hiện tại có giá trị MSE của tập kiểm định lớn hơn giá trị tương ứng ở vòng lặp liền trước.

Việc chương trình ngay lập tức dừng lại khi giá trị MSE của tập kiểm định tăng thông thường không phải là mô hình tốt nhất bởi vì có thể thời điểm đó mô hình đang rơi vào một khu vực hội tụ cục bộ và nếu vượt qua khu vực đó mô hình sẽ được cải thiện và tốt hơn rất nhiều. Do vậy, trong phần lớn các trường hợp chúng ta nên chờ đợi và theo dõi thêm trong một số vòng lặp. Nếu sau thời gian đó, mô hình không được cải thiện thì chương trình mới dừng lại. Trong trường hợp này, công thức lệnh (2) được thêm đối số *patience* và có dạng như sau:

$$es2 = \text{EarlyStopping}(\text{monitor}='val_loss', \text{mode}='min', \text{patience} = 50) \quad (3)$$

trong đó 50 có nghĩa là quá trình huấn luyện sẽ đợi thêm 50 vòng lặp nữa, nếu trong 50 vòng lặp đó mà mô hình không được cải thiện thì chương trình huấn luyện mới chính thức dừng lại. Số vòng lặp được lựa chọn phụ thuộc vào mô hình được xây dựng và đặc điểm của vấn đề đang được nghiên cứu.

Có thể nhận thấy rằng, việc sử dụng đối số *patience* cũng gặp phải một số vấn đề làm giảm hiệu quả của thuật toán. Thứ nhất, số lượng vòng lặp dùng cho đối số *patience* phụ thuộc nhiều vào kinh nghiệm của người sử dụng. Thứ hai, sau số lượng vòng lặp đó thì mô hình thu được lúc này có thể không tốt bằng mô hình tương ứng với thời điểm ngay khi giá trị MSE của tập kiểm định tăng lên. Để tránh những vấn đề này, một phương pháp khác có thể được áp dụng thông qua kỹ thuật nhớ mô hình (Model Checkpoint - MC) được trình bày sau đây.

3.2. Kỹ thuật nhớ mô hình

Kỹ thuật MC xuất phát từ một nguyên tắc đơn giản như sau: sau khi kết thúc một bước lặp mô hình tìm được sẽ được ghi nhớ lại và trong bước lặp tiếp theo nếu mô hình tìm được mới tốt hơn mô hình trước đó thì mô hình mới được ghi lại thay thế cho mô hình cũ. Trong nghiên cứu này, mô hình mới được đánh giá là tốt hơn mô hình cũ nếu như giá trị MSE của tập kiểm định tính theo mô hình mới nhỏ hơn giá trị tính theo mô hình cũ. Rõ ràng, với nguyên tắc này thì mô hình được ghi nhớ cuối cùng là mô hình có giá trị MSE đối với tập kiểm định là nhỏ nhất và được xem là mô hình tốt nhất thu được. Kỹ thuật MC có thể được thực hiện thông qua việc sử dụng lệnh ModelCheckpoint có trong thư viện Keras với cấu trúc như sau:

$$es3 = \text{ModelCheckpoint}('model_best.h5', \text{monitor}='val_loss', \text{mode}='min', \text{save_best_only}=True) \quad (4)$$

trong đó *es3* là tên hàm do người dùng định nghĩa để sử dụng hàm ModelCheckpoint; *model_best.h5* là chỉ ra tên tệp mà mô hình tìm được được ghi lại; *save_best_only=True* là câu lệnh điều khiển là chỉ mô hình tốt nhất tìm được được ghi nhớ và nó sẽ ghi đè lên mô hình tìm được trước đó và các đối số khác tương tự như (3).

So với kỹ thuật ES, kỹ thuật MC rõ ràng đảm bảo mô hình tìm được là mô hình tốt nhất, tuy nhiên số lần lặp của kỹ thuật này lại lớn hơn rất nhiều dẫn đến thời gian huấn luyện dài hơn. Điều này có thể trở thành vấn đề nghiêm trọng đối với các mô hình lớn mà ở đó thời gian huấn luyện là khá dài. Để khắc phục điều này, chúng ta có thể kết hợp hai kỹ thuật ES và MC trong đó ES đóng vai trò dừng chương trình huấn luyện lại khi mà mô hình không có sự tiến bộ đáng kể nào và MC đóng vai trò ghi nhớ lại mô hình tốt nhất tìm được trước khi chương trình huấn luyện dừng lại. Rõ ràng, sự kết hợp này phát huy thế mạnh cũng như khắc phục nhược điểm khi sử dụng riêng từng kỹ thuật ES và MC.

4. Kết quả huấn luyện của các kỹ thuật

Trong phần này chúng ta sẽ xem xét cụ thể hiệu quả của các kỹ thuật đề cập đến ở trên trong bài toán giảm phẳng 39 thanh. Số dữ liệu của tập huấn luyện, tập kiểm định và tập kiểm tra lần lượt là

5000, 2000 và 2000 với điều kiện đảm bảo là không có sự trùng lặp dữ liệu trong các tập đó. Mô hình huấn luyện lấy như trình bày trong phần 2. Số vòng lặp được lấy là 5000. Năm kỹ thuật được nghiên cứu là học sâu thông thường DL tức là số vòng lặp chạy đến giá trị được định trước (5000); kỹ thuật ES thông thường tức là chương trình huấn luyện dừng ngay khi MSE đối với tập kiểm định bị giảm; kỹ thuật ES có *patience* = 500 (ES500); kỹ thuật MC thông thường tức là chương trình huấn luyện chạy đến hết 5000 vòng lặp và mô hình tốt nhất được lưu lại trong quá trình đó; và kết hợp kỹ thuật MC và ES với *patience* = 500 (MC+ES500).

Bảng 1. Kết quả huấn luyện của các kỹ thuật khác nhau

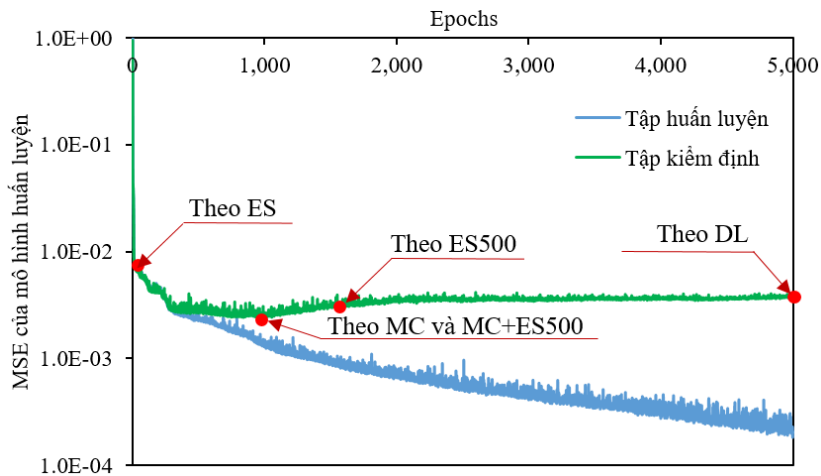
	MSE của mô hình	DL	ES	ES500	MC	MC+ES500
Đối với tập huấn luyện	Nhỏ nhất	1,17E-04	6,30E-03	8,55E-04	1,37E-03	1,41E-03
	Lớn nhất	2,09E-04	8,61E-03	1,06E-03	1,84E-03	1,85E-03
	Trung bình	1,57E-04	7,60E-03	9,59E-04	1,68E-03	1,66E-03
	Độ lệch chuẩn	3,34E-05	7,36E-04	7,46E-05	1,89E-04	1,83E-04
Đối với tập kiểm định	Nhỏ nhất	3,79E-03	5,79E-03	2,84E-03	2,27E-03	2,41E-03
	Lớn nhất	4,43E-03	8,25E-03	4,14E-03	3,01E-03	2,98E-03
	Trung bình	4,02E-03	7,25E-03	3,34E-03	2,57E-03	2,69E-03
	Độ lệch chuẩn	2,57E-04	8,32E-04	5,23E-04	2,71E-04	2,49E-04
Đối với tập kiểm tra	Nhỏ nhất	4,10E-03	5,98E-03	2,78E-03	2,58E-03	2,65E-03
	Lớn nhất	4,76E-03	7,69E-03	4,38E-03	3,01E-03	3,06E-03
	Trung bình	4,60E-03	7,16E-03	3,45E-03	2,80E-03	2,84E-03
	Độ lệch chuẩn	1,89E-03	5,74E-04	1,51E-03	1,16E-03	1,17E-03
Số vòng lặp trung bình		5000	20	1440	5000	1481

Kết quả tổng hợp được trình bày trong Bảng 1, trong đó mỗi kỹ thuật được chạy độc lập 20 lần khác nhau. Dựa vào Bảng 1 ta có thể thấy rằng số lần lặp của DL và MC là 5000 lớn hơn rất nhiều so với của ES500 và MC+ES500 (trung bình khoảng 1400 lần lặp). Số lần lặp trung bình của ES chỉ là 20, nhỏ hơn khá nhiều so với các kỹ thuật khác do điều kiện dừng chương trình huấn luyện của mô hình chỉ là nếu MSE của tập kiểm định ở vòng lặp tiếp theo lớn hơn vòng lặp hiện tại.

Tiếp theo chúng ta đánh giá việc dừng chương trình huấn luyện của các kỹ thuật sẽ ảnh hưởng đến kết quả huấn luyện như thế nào. Trong Bảng 1 cho thấy khi số vòng lặp lên đến 5000 thì mô hình huấn luyện thu được bằng việc sử dụng DL có kết quả MSE trung bình so với tập huấn luyện chỉ là 1,57E-04, nhỏ hơn rất nhiều so với kết quả của các kỹ thuật khác: 7,60E-03, 9,59E-04, 1,68E-03 và 1,66E-03 tương ứng với ES, ES500, MC và MC+ES500. Tuy nhiên, MSE của mô hình thu được khi sử dụng DL đối với tập kiểm tra chỉ là 4,60E-03, không tốt bằng kết quả thu được khi sử dụng ES500, MC và MC+ES500 với các giá trị tương ứng là 3,45E-03, 2,80E-03 và 2,84E-03. Điều này là do hiện tượng mô hình quá khớp đã xảy ra trong quá trình huấn luyện khi sử dụng DL. Tuy nhiên, kết quả thu được khi sử dụng DL lại tốt hơn khi sử dụng ES với MSE trung bình đối với tập kiểm tra của là 7,16E-03. Như vậy ở đây ta có thể thấy rằng kỹ thuật ES đã khiến cho chương trình huấn luyện dừng quá sớm nên mô hình thu được cho kết quả đối với cả tập huấn luyện và tập kiểm tra đều không tốt.

So sánh giữa 3 kỹ thuật ES500, MC và MC+ES500 ta thấy rằng MC và MC+ES500 cho giá trị MSE của mô hình là như nhau và tốt hơn của ES500. Nguyên nhân của vấn đề này là do mô hình huấn luyện thu được từ kỹ thuật MC và MC+ES500 đều là mô hình tốt nhất mà ta có thể tìm được trong tất

cả các bước lặp, trong khi đó kỹ thuật ES500 chỉ tìm ra mô hình huấn luyện tương ứng với thời điểm vòng lặp cách vị trí vòng lặp của mô hình tốt nhất là 500. Bên cạnh đó, mô hình MC+ES500 chỉ sử dụng số vòng lặp trung bình là 1481, nhỏ hơn rất nhiều so với số vòng lặp 5000 của MC. Như vậy, có thể kết luận rằng kỹ thuật MC+ES500 đã cho kết quả mô hình tìm được là tốt nhất. Để mô tả rõ hơn thời điểm thu được các mô hình huấn luyện của các kỹ thuật, một lần huấn luyện tùy ý được thực hiện kết quả được thể hiện trong Hình 4. Dựa vào Hình 4, có thể thấy rằng kỹ thuật ES khiến mô hình dừng khá sớm khi MSE đối với tập huấn luyện và tập kiểm định khá lớn. Đối với kỹ thuật ES500, do mô hình thu được là tại thời điểm sau mô hình tốt nhất 500 vòng lặp nên kết quả MSE đối với tập kiểm định lớn hơn so với khi sử dụng kỹ thuật MC và MC+ES500. Kỹ thuật DL có mô hình huấn luyện thu được tại thời điểm số vòng lặp là 5000 nên có MSE của tập huấn luyện là nhỏ nhất nhưng MSE của tập kiểm định lại lớn hơn khá nhiều so với kết quả của ES500, MC và MC+ES500. Kỹ thuật MC và MC+ES500 tìm được mô hình huấn luyện có giá trị MSE với tập kiểm định là bé nhất.



Hình 4. Thời điểm thu được mô hình huấn luyện của các kỹ thuật

5. Kết luận

Trong bài báo này, hiện tượng mô hình quá khớp đã được trình bày cho bài toán xây dựng mô hình học sâu để ước lượng khả năng chịu tải của giàn phi tuyến. Các kỹ thuật xử lý mô hình quá khớp bao gồm dừng sớm (ES), nhớ mô hình (MC) và kết hợp 2 kỹ thuật này (MC+ES) đã được trình bày và đánh giá sự hiệu quả. Một giàn phẳng gồm 39 thanh được sử dụng để minh họa cho nghiên cứu với tập dữ liệu có thông số đầu vào là diện tích mặt cắt ngang của các thanh giàn và thông số đầu ra là hệ số khả năng chịu tải (LF). Kết quả nghiên cứu cho thấy rằng, xét về cả góc độ thời gian huấn luyện thì ES không sử dụng vòng lặp trễ có thời gian huấn luyện ngắn nhất, còn DL và MC có thời gian huấn luyện dài nhất. Xét dưới góc độ độ chính xác của mô hình, kỹ thuật MC và MC+ES với ES sử dụng kỹ thuật vòng lặp trễ cho kết quả tốt nhất khi mô hình huấn luyện tìm được là mô hình tốt nhất trong trong tất cả các vòng lặp. Từ kết quả nghiên cứu trên, kỹ thuật MC+ES có sử dụng vòng lặp trễ được khuyến nghị áp dụng để xử lý hiện tượng mô hình quá khớp trong quá trình xây dựng mô hình học sâu.

Tài liệu tham khảo

- [1] AISC-LRFD (1999). *Manual of steel construction – load and resistance factor design*. Chicago (IL): American Institute of Steel Construction.
- [2] EN 1993-1-1 (2005). *Eurocode 3: Design of steel structures – part 1-1: general rules and rules for building*. Brussels: European Committee for Standardization.
- [3] Truong, V.-H., Kim, S.-E. (2018). [A robust method for optimization of semi-rigid steel frames subject to seismic loading](#). *Journal of Constructional Steel Research*, 145:184–195.
- [4] Truong, V.-H., Kim, S.-E. (2018). [Reliability-based design optimization of nonlinear inelastic trusses using improved differential evolution algorithm](#). *Advances in Engineering Software*, 121:59–74.
- [5] Ha, M.-H., Vu, Q.-A., Truong, V.-H. (2018). [Optimum design of stay cables of steel cable-stayed bridges using nonlinear inelastic analysis and genetic algorithm](#). *Structures*, Elsevier, 16:288–302.
- [6] Truong, V. H., Kim, S.-E. (2017). [An efficient method for reliability-based design optimization of nonlinear inelastic steel space frames](#). *Structural and Multidisciplinary Optimization*, 56(2):331–351.
- [7] Truong, V.-H., Kim, S.-E. (2017). [An efficient method of system reliability analysis of steel cable-stayed bridges](#). *Advances in Engineering Software*, 114:295–311.
- [8] Truong, V.-H., Nguyen, P.-C., Kim, S.-E. (2017). [An efficient method for optimizing space steel frames with semi-rigid joints using practical advanced analysis and the micro-genetic algorithm](#). *Journal of Constructional Steel Research*, 128:416–427.
- [9] Vu, Q.-V., Truong, V.-H., Papazafeiropoulos, G., Graciano, C., Kim, S.-E. (2019). [Bend-buckling strength of steel plates with multiple longitudinal stiffeners](#). *Journal of Constructional Steel Research*, 158:41–52.
- [10] Zhang, Y., Hu, S., Wu, J., Zhang, Y., Chen, L. (2014). [Multi-objective optimization of double suction centrifugal pump using Kriging metamodels](#). *Advances in Engineering Software*, 74:16–26.
- [11] Breiman, L. (2001). [Random forests](#). *Machine Learning*, 45(1):5–32.
- [12] Friedman, J. H. (2002). [Stochastic gradient boosting](#). *Computational Statistics & Data Analysis*, 38(4): 367–378.
- [13] Safavian, S. R., Landgrebe, D. (1991). [A survey of decision tree classifier methodology](#). *IEEE transactions on systems, man, and cybernetics*, 21(3):660–674.
- [14] Worden, K., Lane, A. J. (2001). [Damage identification using support vector machines](#). *Smart Materials and Structures*, 10(3):540.
- [15] Ha, M.-H., Vu, Q.-V., Truong, V.-H. (2020). [A Deep Learning-Based Procedure for Safety Evaluation of Steel Frames Using Advanced Analysis](#). *CIGOS 2019, Innovation for Sustainable Infrastructure*, Springer, 1137–1142.
- [16] Worden, K., Manson, G. (2006). [The application of machine learning to structural health monitoring](#). *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851):515–537.
- [17] Hung, T. V., Viet, V. Q., Thuat, D. V. (2019). [A deep learning-based procedure for estimation of ultimate load carrying of steel trusses using advanced analysis](#). *Journal of Science and Technology in Civil Engineering (STCE)-NUCE*, 13(3):113–123.
- [18] Thai, H.-T., Kim, S.-E. (2011). [Nonlinear inelastic analysis of space frames](#). *Journal of Constructional Steel Research*, 67(4):585–592.