

Ứng dụng trí tuệ nhân tạo trong phân tích dữ liệu thường địa vật lý từ dữ liệu điện từ

Lê Văn Anh Cường^{1,2,*}, Nguyễn Tiến Hóa³, Vũ Trọng Tấn³,
Đình Quốc Tuấn³, Võ Mạnh Khương³, Vũ Đức Tính³, Trần Hải Nam³,
Nguyễn Văn Thuận^{1,2}, Nguyễn Thị Hồng Hà^{1,2}, Huỳnh Văn Tuấn^{1,2}

¹ Trường Đại học Khoa học Tự nhiên

² Đại học Quốc gia Thành phố Hồ Chí Minh

³ Đoàn Địa vật lý và địa chất biển, Liên đoàn Bản đồ và Địa chất biển miền Nam

* Tác giả liên hệ: lvacuong@hcmus.edu.vn

Ngày gửi bài: 12/6/2025

Ngày chấp nhận đăng:
19/9/2025

Từ khóa: Từ tellua, Ra đa xuyên đất, điện từ tần số cao, điện từ tần số thấp, mạng MLP, Mạng nơ ron tích chập.

Tóm tắt. Dữ liệu trường điện từ đóng vai trò cực kỳ quan trọng trong biểu diễn cấu trúc địa chất và đối tượng dị thường tồn tại từ độ sâu nông đến hàng chục kilomet. Ứng dụng Trí tuệ nhân tạo vào quy trình phân tích dữ liệu địa vật lý này có thể mang đến những lợi ích thiết thực như là, cung cấp kết quả nhanh, chính xác và tiết kiệm nguồn lực của chuyên gia. Trong nội dung nghiên cứu, chúng tôi tập trung vào xây dựng bộ dữ liệu điện từ tần số thấp (từ tellua) hoàn chỉnh bằng việc bổ sung các dữ liệu đo đạc còn thiếu và xác định các đối tượng dị vật ngầm từ dữ liệu điện từ tần số cao (ra đa xuyên đất) theo cơ chế phân xạ. Mô hình trí tuệ tổng quát được xây dựng dựa trên các lớp mạng nơ ron liên kết với nhau, như là MLP và CNN, để thực hiện các công việc xây dựng bộ dữ liệu hoàn chỉnh trường điện từ tại khu vực khoáng sản Olympic Dam, Úc và phân biệt dị vật ngầm tán xạ ra đa xuyên đất tại tỉnh Đồng Nai, Việt Nam.

1. Mở đầu

Trong lĩnh vực Địa Vật lý, hình ảnh của các đối tượng khảo sát không được thể hiện trực tiếp bằng các dữ liệu đo đạc thực địa. Thay vào đó, thông tin hữu ích về các đối tượng đó phải rút ra từ quá trình phân tích và minh giải số liệu. Trong phạm vi nghiên cứu, số liệu điện từ thuộc nhóm tần số cao (liên quan đến hiện tượng phản xạ) và nhóm tần số thấp (liên quan đến hiện tượng cảm ứng) được xử lý và minh giải bằng sự hỗ trợ của công nghệ tính toán hiện đại và các quy trình xử lý truyền thống (Fisher et al., 1996; Le et al., 2011; Daniels, 2004; Kang et al., 2020; Le et al., 2019; Le et al., 2016; LeCun et al., 2015; Mehta et al., 2019; Takam Takougang et al., 2015; Van Nguyen et al., 2017).

Phương pháp Radar xuyên đất (GPR-Ground penetrating radar) được sử dụng rộng

rãi và hiệu quả trong công trình ngầm, môi trường và xây dựng. Phương pháp GPR hoạt động theo nguyên tắc sau: sóng điện từ phát ra từ một anten phát dưới dạng xung, lan truyền trong vật chất với vận tốc được quyết định chủ yếu bởi tính chất điện của vật liệu. Khi sóng lan truyền vào vật chất bên dưới mặt đất, nếu nó gặp các bất đồng nhất hoặc các mặt ranh giới giữa các môi trường có tính chất điện khác nhau, một phần năng lượng sóng sẽ phản xạ hoặc tán xạ trở lại mặt đất trong khi phần năng lượng còn lại tiếp tục di chuyển xuống phía dưới. Sóng phản xạ lại được ghi nhận bởi anten thu và lưu trữ trong bộ nhớ của thiết bị để sử dụng cho việc xử lý và phân tích về sau. Ở dữ liệu đo đạc GPR, sự tán xạ sóng điện từ dưới hình dáng hyperbol đóng vai trò quan trọng để hiểu rõ sự hiện diện của các đối tượng dị vật ngầm

như hệ thống dây/ ống ngầm hoặc các vị trí di thường địa chất (khu vực sụt lún hoặc hang các tơ). Đối với dữ liệu đo đạc lớn sẽ cần đến sự đầu tư công sức và thời gian của chuyên gia khi phân tích. Với khối lượng dữ liệu đo đạc lớn, việc phân tích sẽ đòi hỏi đáng kể công sức và thời gian từ các chuyên gia.

Phương pháp từ tellua (MT) là một phương pháp địa vật lý hiệu quả, thu thập và phân tích sóng điện từ (EM) tự nhiên để xây dựng mô hình cấu trúc độ dẫn điện của môi trường địa chất. Về cơ bản, chúng ta đồng thời đo các sóng trường điện (E) và trường từ (B) theo các phương trục giao góc trên bề mặt, cho phép nghiên cứu độ sâu từ vài chục mét đến hàng trăm kilomet. Nguồn gốc của các trường EM tự nhiên khá đa dạng: các tín hiệu tần số thấp thường đến từ sự tương tác giữa bão mặt trời và từ trường Trái Đất, trong khi các tín hiệu tần số cao lại được tạo ra bởi sấm sét trên toàn cầu. Khi những trường này truyền xuống lòng đất, chúng sẽ tạo ra các dòng điện cảm ứng bên dưới bề mặt. Tại mỗi điểm đo, chúng ta ghi lại các thành phần của trường điện (E_x, E_y) và trường từ (H_x, H_y) theo thời gian. Mối quan hệ giữa hai trường này được biểu diễn bằng trở kháng tensor (Z), theo công thức $E=ZH$. Đại lượng này cung cấp thông tin quan trọng về khả năng dẫn điện (hoặc cách điện) của các môi trường đất đá bên dưới. Tuy nhiên, việc thu thập dữ liệu MT trên thực địa thường gặp nhiều thách thức. Nhiễu từ môi trường (như hoạt động của con người hay đường dây điện) hoặc điều kiện địa hình khó khăn có thể dẫn đến thiếu hụt dữ liệu tại một số điểm đo hoặc các tần số thu thập. Do đó, lấp đầy các khoảng trống dữ liệu này trở nên cực kỳ quan trọng, đảm bảo quá trình phân tích và giải đoán cấu trúc mô hình dẫn điện môi trường được diễn ra chính xác và đáng tin cậy.

Những ứng dụng công nghệ trí tuệ nhân tạo đã được nghiên cứu và sử dụng trong phân tích các dữ liệu địa vật lý (Kang et al., 2020; Le et al., 2025; Perol et al., 2018; Pham and Lefèvre, 2018). Việc ứng dụng trí tuệ nhân tạo có thể giúp các lợi ích sau với hai loại công việc quan trọng: phân tích nhanh và dự đoán các điểm dữ liệu còn thiếu. Trong phạm vi nghiên cứu, các công việc cần thực hiện đối với mỗi quy trình ứng dụng:

(i) Xây dựng bộ dữ liệu: dữ liệu đầu vào và các nhãn thể hiện kết quả phân tích

(ii) Xây dựng cấu trúc mạng nơ ron nhân tạo.

(iii) Thực hiện quá trình huấn luyện dữ liệu và xây dựng mô hình nơ ron nhân tạo

(iv) Thử nghiệm trên các số liệu thực tế đo đạc mới

2. Phương pháp

Ứng dụng các mạng nơ ron sâu và nông được sử dụng để phát hiện các đối tượng dị vật ngầm từ dữ liệu ra đa xuyên đất và dự đoán giá trị đo đạc tại các điểm không gian bị thiếu. Công tác chuẩn bị dữ liệu và thiết lập cấu trúc mạng được thực hiện để xây dựng các mô hình mạng nơ ron này.

Trong nghiên cứu dữ liệu ra đa xuyên đất, đối tượng dị vật ngầm được phát hiện dựa trên các tín hiệu tán xạ hyperbol trên các lát cắt đo đạc thực tế. Các dữ liệu đo đạc thô được xử lý để tạo ra dữ liệu có thể minh giải theo phương pháp truyền thống. Các dữ liệu nhãn sẽ là các kết quả minh giải có hai đối tượng cụ thể: đối tượng thuộc vùng không gian tán xạ hyperbol và đối tượng không thuộc không gian tán xạ hyperbol. Cấu trúc mạng học sâu CNN được sử dụng để học các đặc trưng tán xạ trên các ảnh dữ liệu.

Trong nghiên cứu cấu trúc địa chất sâu, tập dữ liệu trong không gian ba chiều gồm hai chiều không gian theo khoảng cách (trục Tây – Đông và trục Nam – Bắc) và chiều tần số (có liên hệ với trục độ sâu theo mối quan hệ phương trình hiệu ứng lớp da) là các giá trị điện trở suất biểu kiến đo đạc của thành phần tổng trở tellua theo phương không chéo hóa. Đối với nhiệm vụ dự đoán các điểm dữ liệu còn thiếu khi đo đạc, vai trò các điểm đo đạc dữ liệu đầy đủ quan trọng giúp xây dựng mạng nơ ron nông. Do tính chất địa chất khu vực mỗi vùng khác nhau, nên trong phạm vi nghiên cứu, dữ liệu thu thập ở vùng đo đạc Olympic Dam, Úc được sử dụng để dự đoán chính các vị trí đo đạc thiếu dữ liệu cũng tại vùng này.

2.1 Lý thuyết tổng quát xây dựng mô hình mạng nơ ron

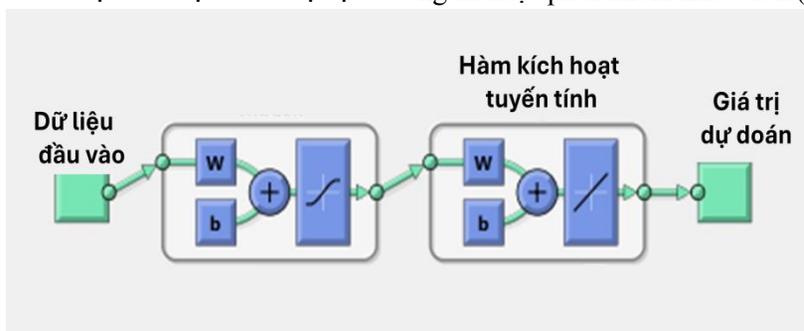
Mạng nơ ron còn được gọi là mạng trí tuệ nhân tạo, phản ánh các hành vi của não người,

cho phép các chương trình máy tính có thể nhận ra các mẫu ảnh và giải quyết nhiều bài toán trong các lĩnh vực trí tuệ nhân tạo, máy học và học sâu (Islam and Kim, 2019; LeCun et al., 2015; Long et al., 2015; Nielsen, 2015). Mạng nơ ron gồm nhiều lớp nốt với các chức năng khác nhau như là, lớp dữ liệu đầu vào, nhiều lớp ẩn và một lớp đầu ra. Mỗi nốt, một điểm nơ ron, nối với các nốt khác và có các trọng số và ngưỡng. Nếu đầu ra của một nốt bất kỳ cao hơn một giá trị ngưỡng thì nốt đó sẽ được kích hoạt và gửi dữ liệu của nó đến lớp tiếp theo của mạng. Nếu bé hơn giá trị ngưỡng, không có dữ liệu của nó được truyền đến lớp tiếp theo của mạng.

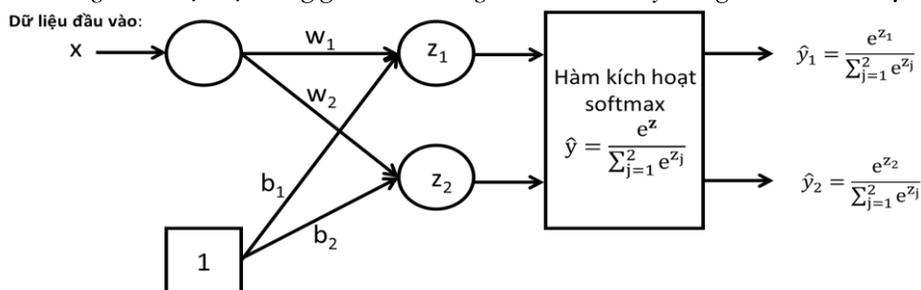
Mạng nhân tạo dựa vào việc “dạy” các dữ liệu để học và tăng độ chính xác của mạng theo thời gian. Một khi các thuật toán học được tinh chỉnh tốt cho độ chính xác, mạng nơ ron là những công cụ cực mạnh trong việc phân loại, dự đoán giá trị và xác định dữ liệu ở tốc độ cực

cao. Các nhiệm vụ như nhận diện giọng nói và nhận diện ảnh có thể được thực hiện trong vài phút so với hàng giờ liền khi được thực hiện bởi các chuyên gia.

Bài toán thuận: Xem mạng nhân tạo đóng vai trò như một hàm bài toán thuận thì các thông tin tọa độ không gian và tần số dữ liệu hoặc ảnh địa vật lý được xem là dữ liệu đầu vào x . Mô hình bài toán thuận bao gồm các giá trị tham số θ (gồm các trọng số “weights” w và giá trị “bias” b). Đối với nhiệm vụ dự đoán các vùng đo đạc bị thiếu dữ liệu, dữ liệu đầu vào sẽ gồm các tọa độ không gian theo hướng Nam-Bắc và Tây-Đông và tần số dữ liệu điện từ được đưa vào bài toán thuận này và cho ra kết quả dự đoán giá trị điện trở biểu kiến của trường điện từ tần số thấp. Còn ở nhiệm vụ xác định dị vật ngầm, các điểm pixel trên ảnh ra đa xuyên đất sẽ được đưa vào bài toán thuận này và cho ra kết quả dự đoán thuộc phân nhóm dị vật tán xạ ngầm hoặc phân nhóm nhiễu nền (background).



Hình 1. Cấu trúc mạng nơ ron đóng vai trò bài toán thuận trong dự đoán giá trị dữ liệu bị thiếu từ dữ liệu đầu vào gồm các tọa độ không gian theo hướng Nam-Bắc và Tây-Đông và tần số dữ liệu điện từ



Nốt $z_1 = w_1 * x + b_1$ liên hệ với nhóm 1 (nhóm hyperbol)

Nốt $z_2 = w_2 * x + b_2$ liên hệ với nhóm 2 (nhóm nhiễu nền)

\hat{y}_1 đại diện cho xác suất khẳng định dữ liệu đầu vào thuộc nhóm 1

\hat{y}_2 đại diện cho xác suất khẳng định dữ liệu đầu vào thuộc nhóm 2

Hình 2. Cấu trúc mạng nơ ron đóng vai trò bài toán thuận trong xác định các vị trí điểm ảnh “pixel” thuộc đối tượng nhóm 1 (nhóm hyperbol tán xạ) hoặc đối tượng nhóm 2 (nhóm nhiễu nền).

Giải bài toán ngược: Việc xác định các trọng số w (w_1, w_2) và giá trị “bias” b (b_1, b_2) được xem như việc giải bài toán ngược. Xây

dựng hàm mất mát “loss function” thích hợp cho từng mục tiêu riêng biệt: bài toán dự đoán và bài toán phân loại. Việc cập nhật các

giá trị tham số θ (gồm các trọng số weights w và giá trị b bias) dựa trên giá trị khởi tạo θ và đạo hàm bậc nhất của hàm mất mát theo θ .

Gọi lớp ẩn \mathbf{z} được hình thành từ dữ liệu \mathbf{x} đầu vào và các trọng số khởi tạo:

$$z_1 = w_1 * x + b_1 \quad (1)$$

$$z_2 = w_2 * x + b_2 \quad (2)$$

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} b_1 & w_1 \\ b_2 & w_2 \end{bmatrix} \begin{bmatrix} 1 \\ x \end{bmatrix} = \boldsymbol{\theta}^T * \mathbf{x};$$

$$\text{Với } \boldsymbol{\theta} = \begin{bmatrix} b_1 & b_2 \\ w_1 & w_2 \end{bmatrix} \text{ và } \mathbf{x} = \begin{bmatrix} 1 \\ x \end{bmatrix}$$

+ Dự đoán các giá trị đo đặc bị thiếu, giá trị điện trở suất biểu kiến được dự đoán ($y_{predict}$) có công thức (MathWorks, 2019b)

$$y_{predict} = \left(\sum_{i=1}^N w_i x + b_i \right) \quad (3)$$

Hàm “mất mát”: $L = \beta (y_{target} - y_{predict})^2 + \alpha (w)^2$ (4)

Hàm mất mát L có hai phần (i) sự sai số của dữ liệu đo đặc thực tế và giá trị dự đoán và (ii) độ trơn của các hệ số w . Các hệ số β và α giúp cân bằng hai yếu tố phần (i) sự sai số của dữ liệu đo đặc thực tế và giá trị dự đoán và (ii) độ trơn của các hệ số w trong quá trình tối thiểu hóa hàm mất mát (Foresee and Hagan, 1997; MacKay, 1992).

+ Phân loại đối tượng dị vật ngầm: Câu trả lời là đúng là dị vật ngầm hoặc là không là dị vật ngầm được thể hiện qua giá trị tính xác suất dữ liệu đầu vào \mathbf{x} thuộc từng nhóm.

Xây dựng hàm mất mát từ đầu ra của hàm kích hoạt softmax và sự phân loại của dữ liệu đầu vào: $L = -\sum_{j=1}^2 y_j * \log \hat{y}_j = -y_1 \log \hat{y}_1 - y_2 \log \hat{y}_2$ (5)

Với, \mathbf{x} thuộc nhóm hyperbol thì $\mathbf{y} = [y_1 \ y_2] = [1 \ 0]$. Hoặc là, \mathbf{x} thuộc nhóm nhiễu nền thì $\mathbf{y} = [y_1 \ y_2] = [0 \ 1]$. Kỹ thuật gán giá trị \mathbf{y} này được gọi là kỹ thuật “one-hot encoding” trong gán sự đặc trưng cho từng loại nhóm dữ liệu (nhóm hyperbol hoặc nhóm nhiễu nền).

Sử dụng hàm kích hoạt (activation function) là hàm toán học softmax để ước tính xác suất dữ liệu đầu vào \mathbf{x} thuộc từng nhóm

$$\hat{y}_1 = \frac{e^{z_1}}{\sum_{j=1}^2 e^{z_j}}, \hat{y}_2 = \frac{e^{z_2}}{\sum_{j=1}^2 e^{z_j}} \quad (6)$$

Tính đạo hàm L'_θ của hàm mất mát theo biến $\boldsymbol{\theta}$ với các công thức sau:

$$L'_\theta = \begin{bmatrix} \frac{\partial L}{\partial b_1} & \frac{\partial L}{\partial b_2} \\ \frac{\partial L}{\partial w_1} & \frac{\partial L}{\partial w_2} \end{bmatrix} \quad (7)$$

Cập nhật tham số mô hình của mạng nơ ron nhân tạo:

$$\boldsymbol{\theta} = \boldsymbol{\theta} - \eta * L'_\theta, \quad (8)$$

với η là hệ số học (trong nghiên cứu thường chọn $\eta = 0.01$).

Việc tính toán đạo hàm L'_θ có thể áp dụng cho từng dữ liệu một hoặc một tập hợp dữ liệu con và hoặc là toàn bộ dữ liệu một lúc. Việc cài đặt số lượng dữ liệu cho một lần học máy được gọi là mini-batch.

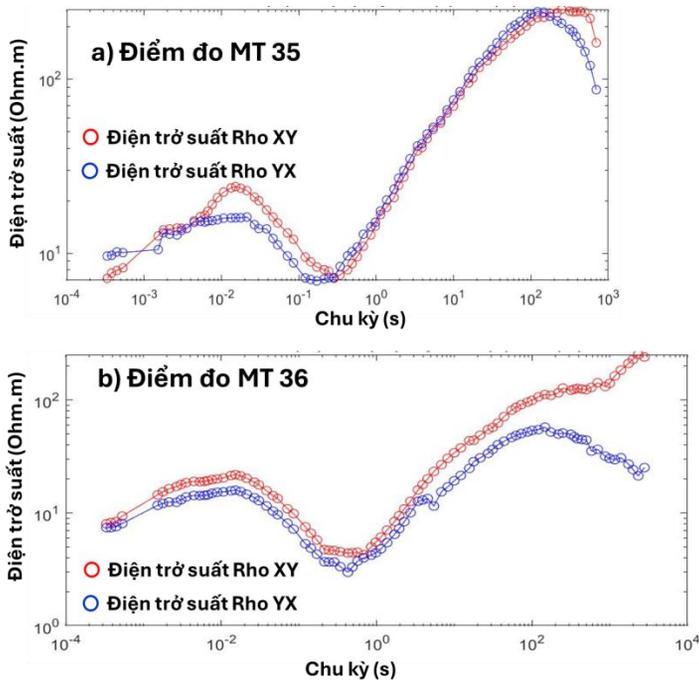
3. Kết quả

3.1. Xây dựng bộ dữ liệu đo đặc điện từ tần số thấp bằng việc dự đoán các vị trí thiếu dữ liệu điểm đo tại khu vực mỏ khoáng sản Olympic Dam

Khu vực khoáng sản Olympic Dam ở Nam Úc nổi tiếng có trữ lượng lớn quặng Nd, vàng, đồng (Haynes et al., 1995; Heinson et al., 2006). Những thành phần quặng kim loại có nguồn gốc từ các lớp vỏ trái đất sâu và manti. Những ranh giới trầm tích được xác định là nhóm Neoproterozoic Wilpena, nằm trên thành tạo Mesoproterozoic Pandurra (Roberts and Hudson, 1983). Các nghiên cứu địa vật lý (như là địa chấn và từ tellua) đã được tiến hành để tìm hiểu cấu trúc địa chất tại khu vực này (Geoscience Australia, 2003, 2018). Vị trí đo đặc từ tellua (MT) gần với các quặng khoáng sản Olympic Dam và bên dưới là lớp phủ Adelaide Rift Complex (Drummond et al., 2006; Heinson et al., 2018; Heinson et al., 2006) (Hình 3). Dữ liệu đo đặc từ tellua đo đặc có 87 chu kỳ từ 0,0003 giây đến 2796,2030 giây. Một ví dụ về sự khác biệt thông tin dữ liệu ở điểm đo MT 35 và 36 (hình 4). Chúng tôi nhận thấy có sự khác biệt đáng kể về số lượng điểm dữ liệu giữa hai điểm đo từ trường (MT) 35 và 36. Cụ thể, tại điểm đo MT 36, chúng tôi thu được nhiều dữ liệu hơn đáng kể, đặc biệt là ở dải tần số thấp với chu kỳ lớn hơn 1000 giây.



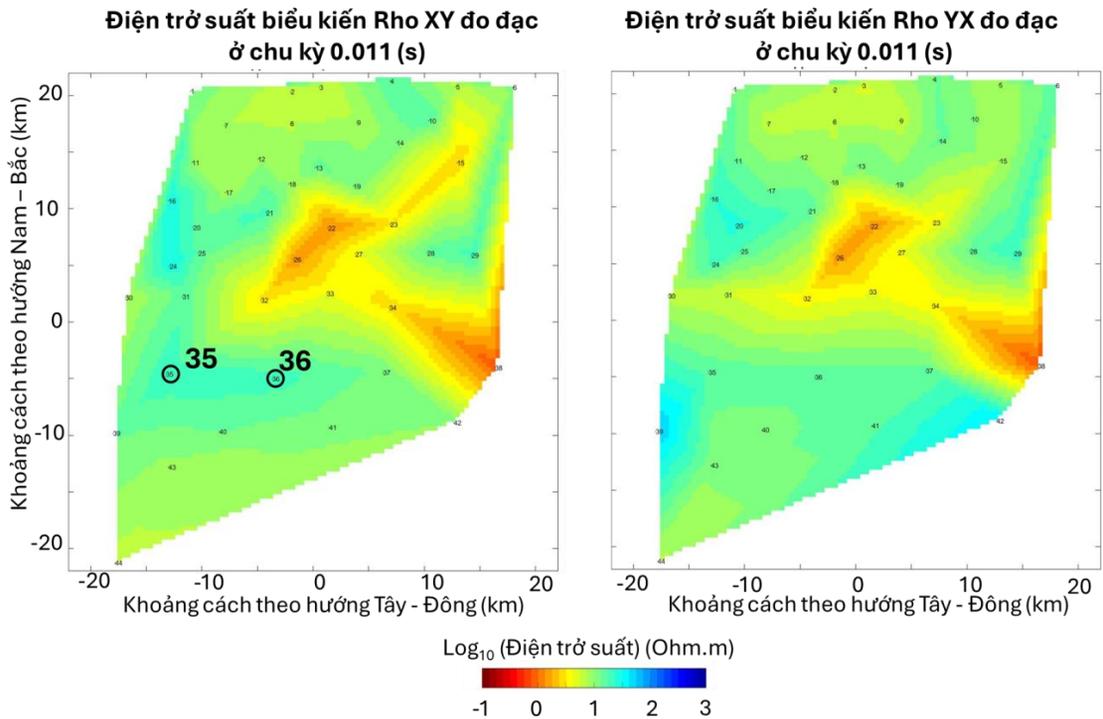
Hình 3. Vị trí đo đặc từ tellua ở khu vực vùng mỏ Olympic Dam, Nam Úc



Hình 4. Dữ liệu đo đặc MT ở hai điểm đo 35 và 36 có sự không giống nhau về số lượng thông tin điểm dữ liệu. Ở điểm đo MT 36, nhiều dữ liệu hơn ở phần chu kỳ lớn hơn 1000 s.

Trong nhiệm vụ dự đoán dữ liệu còn thiếu, để thiết lập mô hình mạng neural, chúng tôi sử dụng gần 1000000 điểm dữ liệu. Trong đó, 70% lượng dữ liệu dùng để tạo mô hình nơ ron, 15% lượng dữ liệu dùng để thẩm định mô hình và 15% lượng dữ liệu dùng để

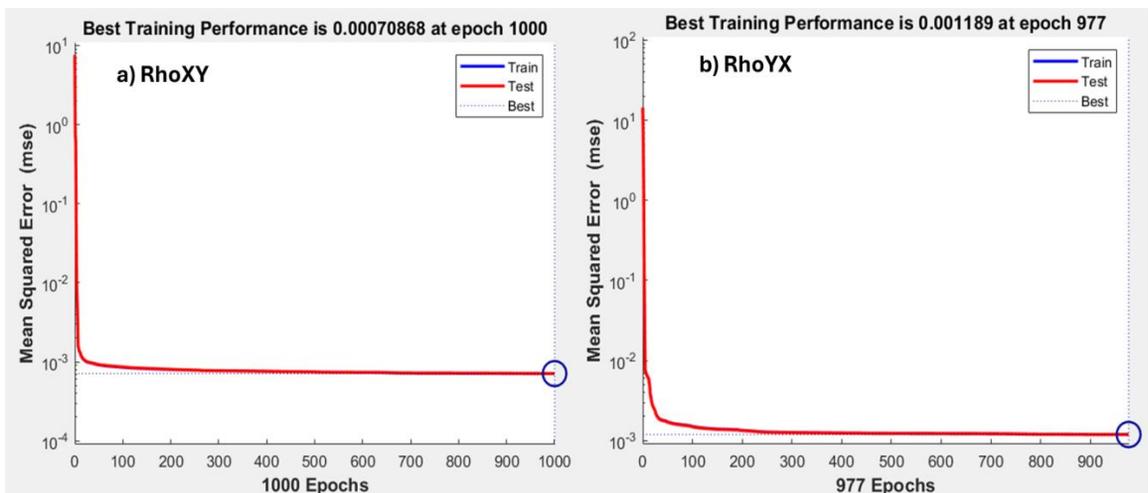
kiểm tra (dữ liệu mà mô hình nơ ron chưa được thấy trong quá trình huấn luyện). Chúng tôi sử dụng dữ liệu điện trở suất của hai thành phần Z_{xy} và Z_{yx} của ten xơ trở kháng để nghiên cứu.



Hình 5. Dữ liệu đo đạc điện trở suất biểu kiến ở chu kỳ 0,011 (s).

Sau khi huấn luyện để xây dựng mô hình neuron, hệ số sai số “mean square error” đều suy giảm rõ rệt và phản ánh sự suy giảm của hàm

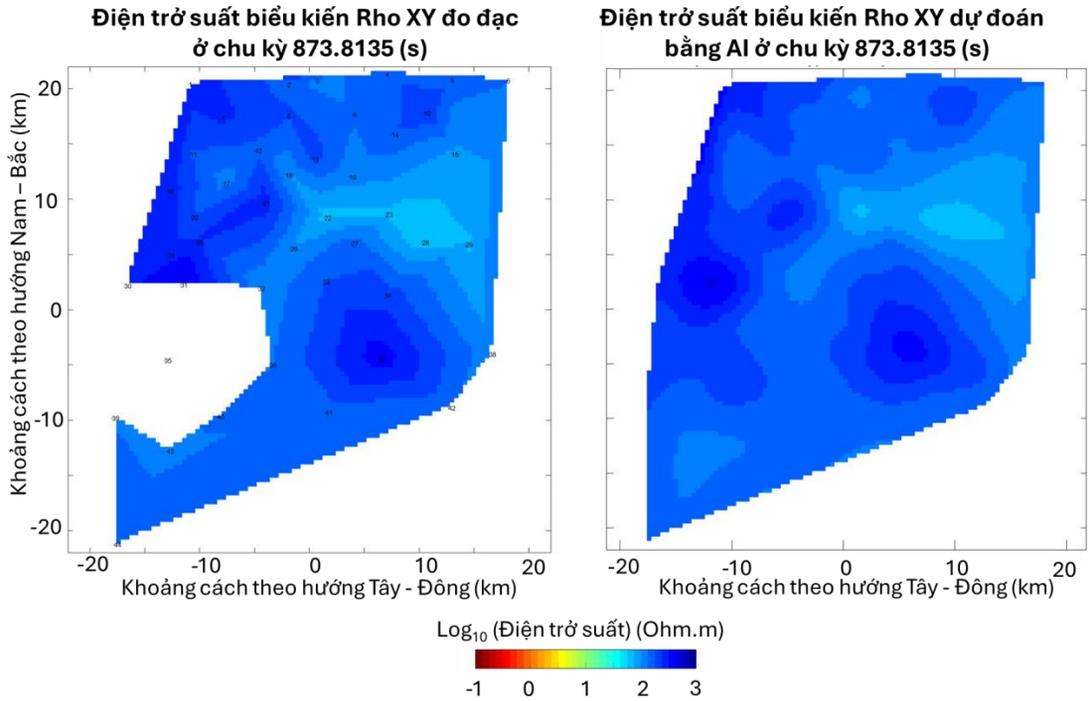
mất mát của cả hai loại dữ liệu điện trở suất biểu kiến Z_{xy} và Z_{yx} (Xem hình 6).



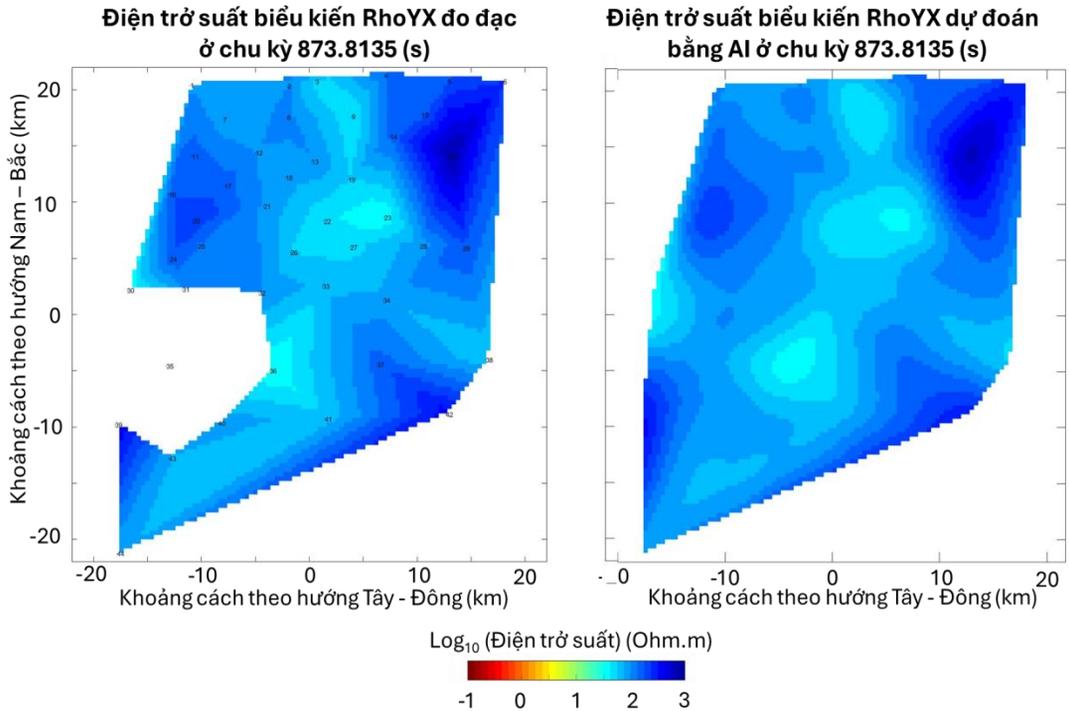
Hình 6. Sai số Mean Squared Error phản ánh hàm mất mát suy giảm của hai dữ liệu điện trở biểu kiến RhoXY và RhoYX

Sử dụng mô hình đã đào tạo được để dự đoán các dữ liệu còn thiếu (xem hình 7 và hình 8). Nhìn chung, các cấu trúc điện trở biểu kiến đo đạc và mô hình hóa thể hiện sự

giống nhau và cho thấy thấy rằng hướng tiếp cận này có thể được nghiên cứu tiếp tục trong việc tạo các dữ liệu mới.



Hình 7. Dữ liệu đo đạc và dự đoán của điện trở suất biểu kiến của Zxy ở chu kỳ 873,8135 (s)



Hình 8. Dữ liệu đo đạc và dự đoán của điện trở suất biểu kiến của Zyx ở chu kỳ 873,8135 (s).

3.2. Kết quả Phát hiện dị vật ngầm điện từ tần số thấp bằng mạng nơ ron tích chập

Mạng nơ ron tích chập là một thuật toán mạnh trong lĩnh vực học sâu (Amidi and

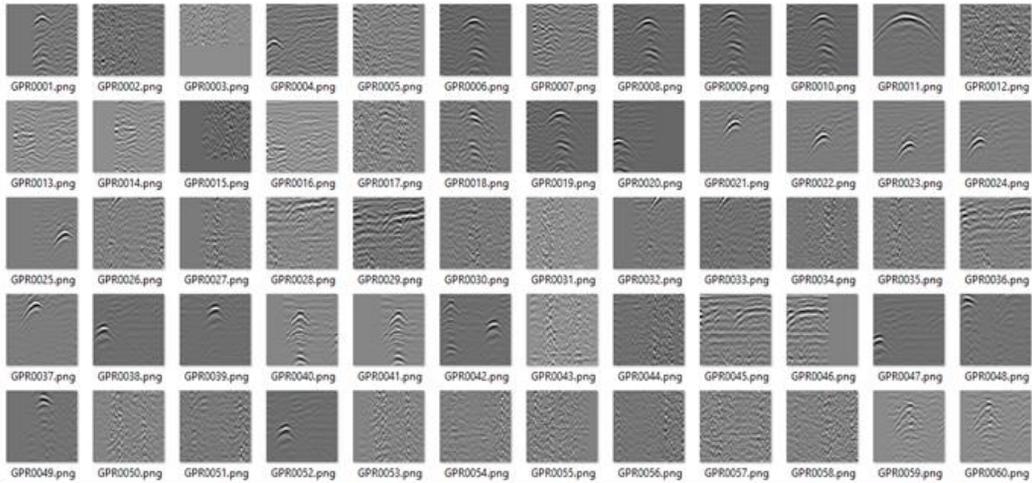
Amidi, 2018; David and Netanyahu, 2016; Stanford, 2023). Mạng sử dụng dữ liệu đầu vào là các hình ảnh, sử dụng các trọng số và độ lệch (theo lý thuyết của mạng trí tuệ nhân tạo) cho

các đặc trưng/ đặc điểm khác nhau trong hình ảnh và có thể phân biệt từng đặc trưng/ đặc điểm này với nhau (Islam and Kim, 2019; LeCun et al., 2015; Long et al., 2015; Nielsen, 2015).

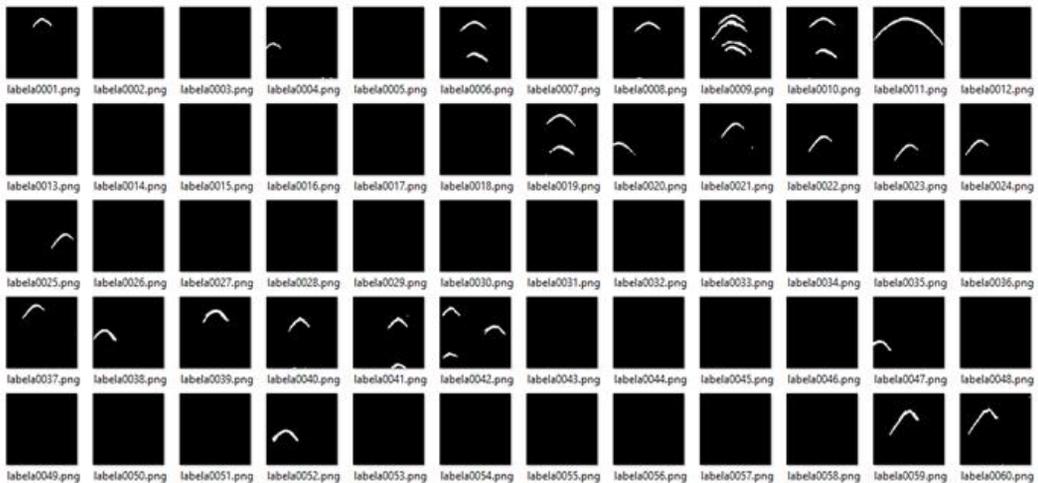
Phương pháp xử lý “Semantic Segmentation” (phân đoạn ngữ nghĩa) được

ngiên cứu để xử lý các tập số liệu điện tử. Phương pháp này bao gồm một mạng lưới nơ ron tích chập với hai hệ thống độc lập “encoder” (mã hóa) và “decoder” (giải mã) (Islam and Kim, 2019) phục vụ cho việc nhận diện các đối tượng di vật ngầm theo các điểm ảnh “pixel” trong số liệu điện tử GPR.

Dữ liệu ảnh đầu vào



Nhãn đầu vào

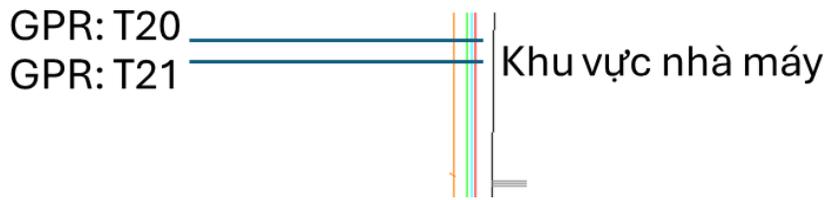


Hình 9. Một số dữ liệu ảnh và nhãn đầu vào phục vụ cho việc học máy: tạo mô hình nơ ron tích chập semantic segmentation

Chúng tôi xây dựng bộ dữ liệu 1000 ảnh và 1000 nhãn (label) (Xem hình 9). Sau đó, chúng tôi thử kết cấu mạng với 600 dữ liệu và 400 dữ liệu test (Xem Hình 9). Sau đó chúng tôi xây dựng mô hình tốt nhất cho cả 1000 ảnh. Bước “training” được tiến hành với việc sử dụng hàm trainNetWork.m

(MathWorks, 2019b). Mô hình kết quả cho độ chính xác trên 91%.

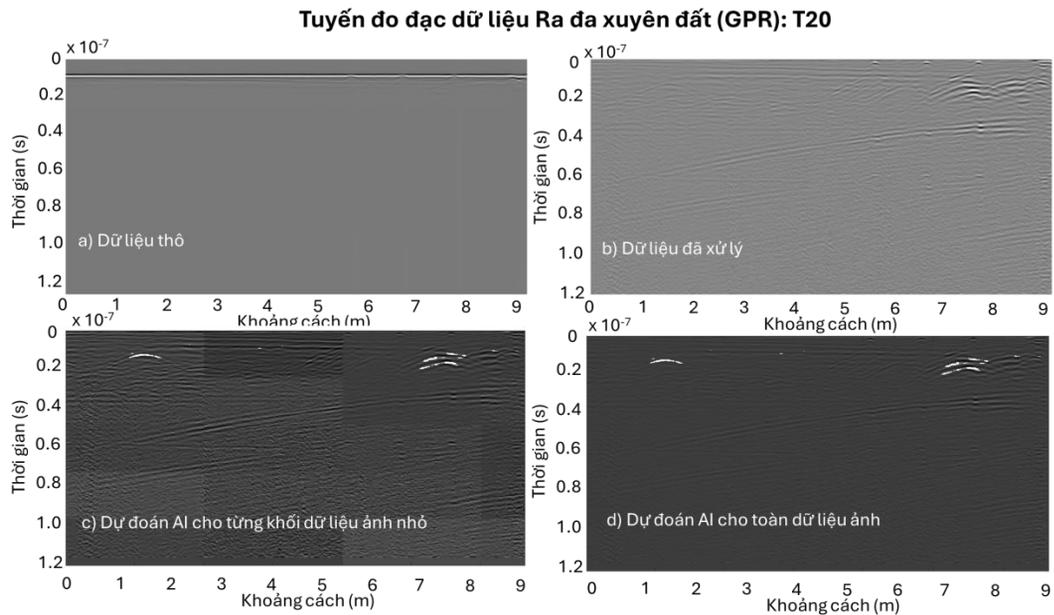
Dữ liệu thử nghiệm ứng dụng mô hình neural được đo đạc tại một nhà máy tại Khu công nghiệp, Nhơn Trạch, Tỉnh Đồng Nai (Hình 10). Thiết bị ra đa xuyên đất là máy Duo Detector 700 MHz.



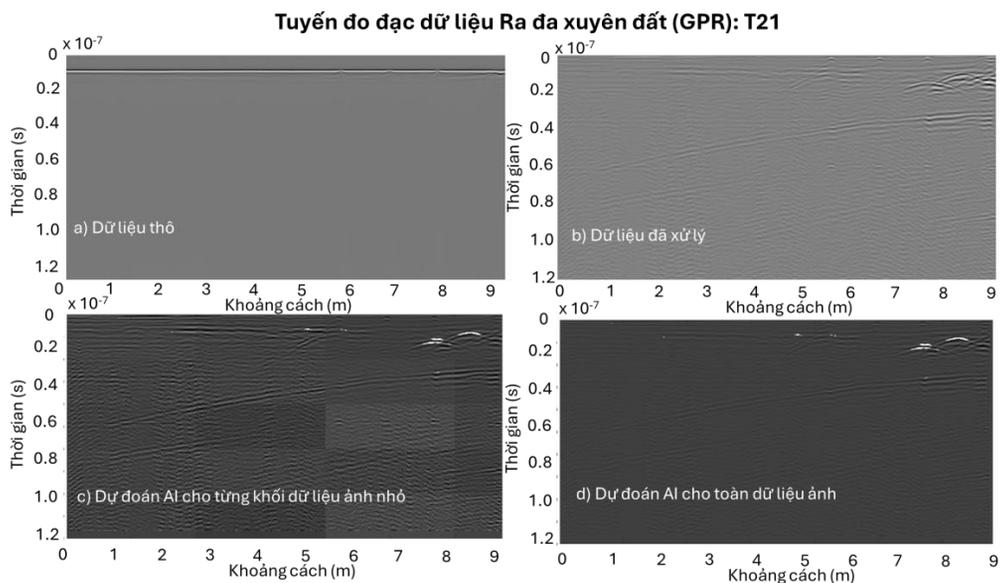
Hình 10. Sơ đồ tuyến đo tại một nhà máy tại Khu công nghiệp, Nhơn Trạch, Tỉnh Đồng Nai

Kết quả xử lý mạng nơ ron (hai hình bên dưới) của tuyến Ra đa xuyên đất T20 và T21 cùng phát hiện ra nhóm đối tượng ở vị trí $x = 7 - 9$ m (Xem hình 11 và hình 12). Tuy

nhiên, thực tế có 3 dị vật ở khu vực này nhưng chương trình nơ ron chỉ phát hiện được 2 đối tượng.



Hình 11. Kết quả xử lý dữ liệu thô và kết quả minh giải bằng ứng dụng nơ ron của tuyến T20



Hình 12. Kết quả xử lý dữ liệu thô và kết quả minh giải bằng ứng dụng nơ ron của tuyến T21

4. Kết luận

Dữ liệu trường điện từ đóng vai trò thiết yếu trong việc biểu diễn cấu trúc địa chất và xác định các dị thường từ nông đến sâu hàng chục kilômét. Việc ứng dụng Trí tuệ nhân tạo (AI) vào quy trình phân tích dữ liệu địa vật lý này mang lại nhiều lợi ích vượt trội, bao gồm: cung cấp kết quả nhanh chóng, nâng cao độ chính xác và tối ưu hóa nguồn lực chuyên gia. Trong nghiên cứu này, chúng tôi đã tập trung cập nhật một bộ dữ liệu điện từ tần số thấp có thành phần điện trở biểu kiến RhoXY và RhoYX hoàn chỉnh thông qua việc bổ sung các dữ liệu còn thiếu. Đồng thời, chúng tôi cũng tiến hành xác định các đối tượng dị vật ngầm dựa trên dữ liệu điện từ tần số cao (radar xuyên đất – GPR) theo cơ chế phản xạ. Mô hình trí tuệ tổng quát được phát triển dựa trên sự kết hợp của các lớp mạng nơ-ron, bao gồm Mạng Đa Lớp (MLP) và Mạng Tích Chập (CNN). Mô hình này được thiết kế để thực hiện hai nhiệm vụ chính: (i) Xây dựng bộ dữ liệu trường điện từ hoàn chỉnh tại khu vực khoáng sản Olympic Dam, Úc và (ii) Phân biệt các dị vật ngầm tán xạ radar xuyên đất tại tỉnh Đồng Nai, Việt Nam.

Lời cảm ơn

Nghiên cứu được tài trợ bởi Đại học Quốc gia Thành phố Hồ Chí Minh (ĐHQG-HCM) trong khuôn khổ đề tài mã số C2024-18-06. Chúng tôi xin gửi lời cảm ơn đến Tổ chức Geoscience Australia, Úc đã hỗ trợ dữ liệu đo đạc và Đại học Curtin, Úc trong hỗ trợ phần mềm tính toán.

Đóng góp: L.V.A.C. đã xử lý và chạy chương trình neural networks. Tất cả các tác giả đóng góp cho phân tích dữ liệu và viết bản thảo. L.V.A.C. chủ yếu viết bản thảo.

Tài liệu tham khảo

Amidi, A., Amidi, S., 2018. Convolutional Neural Networks cheatsheet.

Daniels, D.J., 2004. Ground penetrating radar. *Wiley Online Library*.

David, O.E., Netanyahu, N.S., 2016. DeepPainter: Painter classification using deep convolutional autoencoders, *Artificial Neural Networks and Machine Learning–ICANN 2016: 25th International Conference on Artificial Neural Networks, Barcelona, Spain, September 6-*

9, 2016, Proceedings, Part II 25. Springer, pp. 20-28.

Fisher, S.C., Stewart, R.R., Jol, H.M., 1996. Ground penetrating radar (GPR) data enhancement using seismic techniques. *Journal of environmental and engineering geophysics 1*, 89-96.

Foresee, F.D., Hagan, M.T., 1997. Gauss-Newton approximation to Bayesian learning, *Proceedings of international conference on neural networks (ICNN'97). IEEE, pp. 1930-1935.*

Islam, M., Kim, J.-M., 2019. Vision-based autonomous crack detection of concrete structures using a fully convolutional encoder–decoder network. *Sensors 19*, 4251.

Kang, M.-S., Kim, N., Lee, J.J., An, Y.-K., 2020. Deep learning-based automated underground cavity detection using three-dimensional ground penetrating radar. *Structural Health Monitoring 19*, 173-185.

Le, C.V.A., Harris, B.D., Pethick, A.M., 2019. New perspectives on Solid Earth Geology from Seismic Texture to Cooperative Inversion. *Scientific Reports 9*, 14737.

Le, C.V.A., Harris, B.D., Pethick, A.M., Takam Takougang, E.M., Howe, B., 2016. Semiautomatic and Automatic Cooperative Inversion of Seismic and Magnetotelluric Data. *Surveys in Geophysics 37*, 845-896.

Le, C.V.A., Huynh, T.V., Doan, T.T.T., Nguyen, T.V., 2025. 3D Prior Electrical Resistivity Model Building for Magnetotelluric Inversion Using Neural Network Algorithm to the Kevitsa Mineral Deposit, Finland. *IEEJ TRANSACTIONS ON ELECTRICAL AND ELECTRONIC ENGINEERING*.

Le, C.V.A., Nguyen, V.T., Nguyen, G.V., Dang, H.T., Vo, T.M., 2011. Nghiên cứu dịch chuyển Kirchoff để xác định mô hình vận tốc trong phương pháp điện từ tần số cao. *Tạp chí các khoa học về Trái đất 33*, 142-150.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *Nature 521*, 436-444.

Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation, *Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3431-3440.*

MacKay, D.J., 1992. Bayesian interpolation. *Neural computation 4*, 415-447.

MathWorks, 2019a. Bayesian regularization backpropagation, 2014b ed.

MathWorks, 2019b. Train a neural network, R2019a ed.

Mehta, S., Paunwala, C., Vaidya, B., 2019. CNN based traffic sign classification using Adam optimizer, 2019 international conference on intelligent computing and control systems (ICCS). *IEEE*, pp. 1293-1298.

Nielsen, M., 2015. Neural Networks and Deep Learning.

Perol, T., Gharbi, M., Denolle, M., 2018. Convolutional neural network for earthquake detection and location. *Science Advances* 4, e1700578.

Pham, M.-T., Lefèvre, S., 2018. Buried object detection from B-scan ground penetrating radar data using Faster-RCNN, IGARSS 2018-

2018 IEEE international geoscience and remote sensing symposium. *IEEE*, pp. 6804-6807.

Stanford, 2023. Convolutional Neural Networks (CNNs / ConvNets).

Takam Takougang, E., Harris, B., Kopic, A., Le, C.V.A., 2015. Cooperative joint inversion of 3D seismic and magnetotelluric data: With application in a mineral province. *Geophysics* 80, 1-13.

Van Nguyen, T., Van Le, C.A., Nguyen, V.T., Dang, T.H., Vo, T.M., Vo, L.N.N., 2017. Energy Analysis in Semiautomatic and Automatic Velocity Estimation for Ground Penetrating Radar Data in Urban Areas: Case Study in Ho Chi Minh City, Vietnam, *International Conference on Geospatial Technologies and Earth Resources*. Springer, pp. 34-51.

Summary

Application of artificial intelligence in geophysical anomaly analysis from electromagnetic data

Lê Văn Anh Cường^{1,2,*}, Nguyễn Tiến Hóa³, Vũ Trọng Tấn³, Đinh Quốc Tuấn³, Võ Mạnh Khương³, Vũ Đức Tinh³, Trần Hải Nam³, Nguyễn Văn Thuận^{1,2}, Nguyễn Thị Hồng Hà^{1,2}, Huỳnh Văn Tuấn^{1,2}

¹ *University of Science, Ho Chi Minh City, Vietnam*

² *Vietnam National University, Ho Chi Minh City, Vietnam*

³ *South Viet Nam Division of Mapping and Marine Geology, Ho Chi Minh City, Vietnam*

* *Corresponding author: lvacuong@hcmus.edu.vn*

“Electromagnetic field data plays an extremely important role in representing geological structures and anomalous objects that exist from shallow depths to several tens of kilometers. The application of Artificial Intelligence (AI) in the analysis workflow of such geophysical data can bring practical benefits, such as providing fast and accurate results while saving expert resources. In this study, we focus on building a complete low-frequency electromagnetic (magnetotelluric) dataset by supplementing missing survey data and identifying underground anomalous objects from high-frequency electromagnetic data (Ground Penetrating Radar) based on diffraction mechanisms. The general AI model is developed based on interconnected neural network layers, including different networks such as MLP and CNN, to implement tasks of constructing a complete electromagnetic field dataset for the Olympic Dam mineral area in Australia, and distinguishing subsurface scattering objects from Ground Penetrating Radar data in Dong Nai province, Vietnam.”

Keywords: *Magnetotelluric, Ground Penetrating Radar, High Frequency Electromagnetics, Low Frequency Electromagnetics, MLP Network, Convolutional Neural Network.*