

THUẬT TOÁN CHI TIẾT PHÂN LOẠI ẢNH VIỄN THÁM BẰNG MẠNG NƠ RON NHÂN TẠO NHIỀU LỚP

NGUYỄN DƯ KHANG, NGUYỄN THỊ THANH BÌNH

Cục Viễn thám quốc gia

Tóm tắt:

Với sự phát triển nhanh chóng của các phương pháp thu nhận và lưu trữ dữ liệu viễn thám, ngày nay, khối lượng dữ liệu khổng lồ cần được xử lý và phân tích. Các dữ liệu này quá lớn nên khả năng của các chuyên gia không còn đủ nữa, dẫn đến nhu cầu về các phương pháp phân tích dữ liệu tự động liên tục tăng. Ứng dụng mạng nơ ron nhân tạo là một trong những giải pháp giải quyết vấn đề trên. Bài báo đề xuất thuật toán chi tiết huấn luyện mạng nơ ron nhân tạo nhiều lớp và ứng dụng phân loại ảnh viễn thám. Kết quả thực nghiệm, đánh giá độ chính xác cho thấy, phương pháp này hoàn toàn khả thi và đơn giản để bán tự động hóa phân loại các đối tượng thể hiện trên ảnh viễn thám, phục vụ cho việc thành lập, cập nhật dữ liệu không gian địa lý quốc gia.

1. Đặt vấn đề

Ngày nay, có rất nhiều tài liệu về Viễn thám được xuất bản trong và nước ngoài liên quan đến vấn đề tự động hóa phân loại ảnh viễn thám, song, chủ yếu trong các lĩnh vực chuyên đề. Các công trình liên quan đến ứng dụng công nghệ viễn thám để cập nhật dữ liệu không gian địa lý theo hướng bán tự động hóa hầu như thiếu vắng. Những kiến thức cơ bản về mạng nơ ron nhân tạo trình bày trong Tài liệu hướng dẫn “Xử lý và giải đoán ảnh số viễn thám” của các tác giả (sản phẩm đề tài mã số TNMT.2018.08.08). Sau đây trình bày thuật toán chi tiết huấn luyện mạng nơ ron nhân tạo nhiều lớp và ứng dụng phân loại ảnh viễn thám, chưa thấy công bố trong các tài liệu trong và ngoài nước.

2. Cấu trúc mạng

Hình 1 thể hiện cấu trúc mạng nơ ron 3 lớp (hai lớp ẩn và lớp ra). Đối với mạng nơ ron nhiều lớp khác, việc nghiên cứu hoàn toàn tương tự.

Ký hiệu lớp vào của mạng là X, các lớp ẩn là A, B và lớp ra là Y, với số nơ ron (không kể nơ ron chuyển vị) tương ứng là n , na , nb và m .

(Xem hình 1)

Ký hiệu các ma trận trọng số giữa các lớp:

$$\left. \begin{aligned} \text{Lớp vào X và lớp ẩn A: } W^A &= (W^{a_1}, W^{a_2}, \dots, W^{a_{na}})^T \\ W^{a_i} &= (w^{A_{1i}}, w^{A_{2i}}, \dots, w^{A_{(n+1)i}}), \quad i = 1 \div na \\ \text{Lớp ẩn A và lớp ẩn B: } W^B &= (W^{b_1}, W^{b_2}, \dots, W^{b_{nb}})^T \\ W^{b_i} &= (w^{B_{1i}}, w^{B_{2i}}, \dots, w^{B_{(n+1)i}}), \quad i = 1 \div nb \\ \text{Lớp ẩn B và lớp ra Y: } W^Y &= (W^{y_1}, W^{y_2}, \dots, W^{y_m})^T \\ W^{y_i} &= (w^{Y_{1i}}, w^{Y_{2i}}, \dots, w^{Y_{(nb+1)i}}), \quad i = 1 \div m \end{aligned} \right\} \quad (1)$$

Ngày nhận bài: 05/3/2021, ngày chuyển phân biên: 09/3/2021, ngày chấp nhận phân biên: 15/3/2021, ngày chấp nhận đăng: 18/3/2021

Tập hợp các ma trận trọng số của mạng là $W_s = (W^A, W^B, W^Y)$.

Véc tơ đầu vào và đầu ra của các lớp:

Ký hiệu I^A và O^A là véc tơ đầu vào và đầu ra của lớp A, I_i^a và O_i^a ($i = 1 \div na$) là các phần tử của chúng. Tương tự sử dụng ký hiệu đối với các lớp khác.

$$\left. \begin{aligned}
 \text{Lớp vào X: } I^X = O^X &= (x_1, x_2, \dots, x_n, 1)^T \\
 \text{Lớp ẩn A: } I^A &= (I^a_1, I^a_2, \dots, I^a_{na})^T, \quad I^a_i = (W^a_i) \cdot O^X \\
 &O^A = (O^a_1, O^a_2, \dots, O^a_{na}), \quad O^a_i = F(I^a_i), \quad i = 1 \div na \\
 \text{Lớp ẩn B: } I^B &= (I^b_1, I^b_2, \dots, I^b_{nb}), \quad I^b_i = (W^b_i) \cdot O^A \\
 &O^B = (O^b_1, O^b_2, \dots, O^b_{nb}), \quad O^b_i = F(I^b_i), \quad i = 1 \div nb \\
 \text{Lớp ra Y: } I^Y &= (I^y_1, I^y_2, \dots, I^y_m), \quad I^y_i = (W^y_i) \cdot O^B \\
 &O^Y = (O^y_1, O^y_2, \dots, O^y_m), \quad O^y_i = F(I^y_i), \quad i = 1 \div m
 \end{aligned} \right\} \quad (2)$$

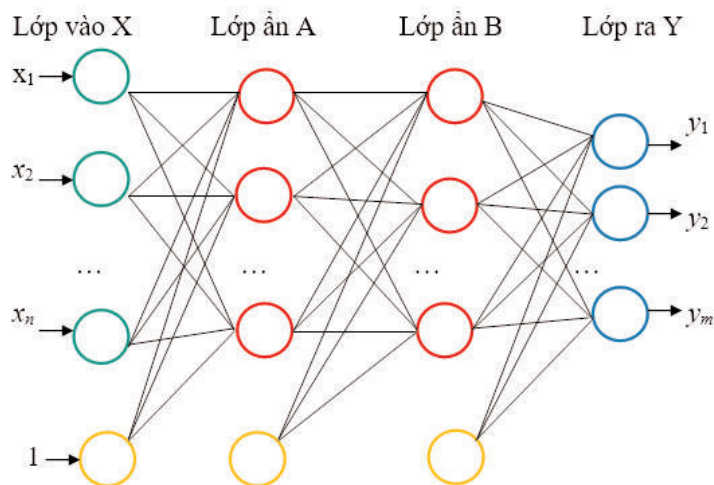
Ở đây F là hàm kích hoạt. Trường hợp mạng chỉ liên quan đến số dương (+), F lấy hàm sigmoid: $F(x) = 1/(1 + e^{-x})$, trường hợp mạng liên quan đến cả dấu âm (-) và dấu dương, F lấy hàm tang hyperbol: $(e^{2x} - 1)/(e^{2x} + 1)$.

3. Thuật toán huấn luyện mạng perceptron 3 lớp

Giả sử cho cặp mẫu (X_s, Y_s) gồm véc tơ đầu vào $X_s = (x_1, x_2, \dots, x_n, 1)^T$ và véc tơ đầu ra $Y_s = (y_1, y_2, \dots, y_m)^T$, trong đó, n - là số tín hiệu đầu vào, m - là số nơ ron của lớp ra.

Huấn luyện mạng là quá trình xác định tập hợp các ma trận trọng số W_s , sao cho véc tơ đầu vào X_s xác định véc tơ đầu ra Y_s ’ so với Y_s thỏa mãn trong hạn sai cho phép.

Các bước huấn luyện mạng (cho một cặp mẫu) tiến hành như sau:



Hình 1: Mạng nơ ron nhiều lớp

Bước 1: Xác định dữ liệu ban đầu

- Số nơ ron (không kể nơ ron chuyển vị) của lớp vào X, các lớp ẩn A, B và lớp ra Y, tương ứng là n, na, nb và m ;

- Véc tơ đầu vào $X_s = (x_1, x_2, \dots, x_n, 1)^T$ và véc tơ đầu ra $Y_s = (y_1, y_2, \dots, y_m)^T$;

- Tốc độ huấn luyện α ; moment β ; giới hạn sai số q và giới hạn vòng lặp *Epoch*;

- Tập hợp các ma trận trọng số W_s lấy ngẫu nhiên trong khoảng $[-0,5; 0,5]$.

Lấy $l = 0$.

Bước 2: Tính $l = l + 1$, nếu $l > Epoch$, chuyển đến bước 6.

- Xác định các véc tơ đầu vào và đầu ra của các lớp: I^A, O^A, I^B, O^B và I^Y, O^Y theo các công thức (1) và (2).

- Tính sai số theo công thức:

$$MSE = \left(\sum_{i=1}^m (y_i - O^Y_i)^2 \right) / m, \text{ nếu } MSE \leq \theta, \text{ chuyển đến bước 6.}$$

Bước 3: Hiệu chỉnh ma trận trọng số W^Y giữa lớp ẩn B và lớp ra Y:

$$W^Y = (W^Y_1, W^Y_2, \dots, W^Y_m)^T,$$

$$W^Y_i = (w^Y_{1i}, w^Y_{2i}, \dots, w^Y_{(nb+1)i}), \quad i = 1 \div m.$$

Lưu ý, trọng số đầu ra của các nơ ron chuyển vị không hiệu chỉnh.

- Xác định véc tơ delta: $\delta^Y = (\delta^Y_1, \delta^Y_2, \dots, \delta^Y_m)$, theo công thức:

$$\delta^Y_i = (O^Y_i - y_i) F'(I^Y_i) = (y_i - O^Y_i) \cdot (1 - O^Y_i) \cdot O^Y_i, \quad i = 1 \div m, \quad (3)$$

$F'(I^Y_i)$ - là đạo hàm của hàm kích hoạt F.

- Xác định gradient của các khớp nối giữa lớp ẩn B và lớp ra Y, theo công thức:

$$\text{grad}^{B_{ji}} = \delta^Y_i \cdot O^{B_{ji}}, \quad j = 1 \div nb, \quad i = 1 \div m. \quad (4)$$

- Xác định gia số hiệu chỉnh:

$$\Delta w^{Y_{ji}} = -\alpha \cdot \text{grad}^{B_{ji}} + \beta \cdot \Delta w^{Y_{ji}}(l-1), \quad j = 1 \div nb, \quad i = 1 \div m, \quad (5)$$

trong đó, $\Delta w^{Y_{ji}}(l-1)$ - là gia số hiệu chỉnh của vòng lặp trước.

- Hiệu chỉnh các véc tơ trọng số:

$$w^{Y_{ji}} = w^{Y_{ji}} + \Delta w^{Y_{ji}}, \quad j = 1 \div nb, \quad i = 1 \div m. \quad (6)$$

Bước 4: Hiệu chỉnh ma trận trọng số giữa các lớp ẩn A và B:

- Xác định véc tơ delta của lớp ẩn B: $\delta^B = (\delta^{B_1}, \delta^{B_2}, \dots, \delta^{B_{nb}})$, theo công thức:

$$\delta^{B_i} = (1 - O^{B_i}) \cdot O^{B_i} \cdot \sum_{j=1}^m (w^{Y_{ij}} \cdot \delta^Y_j), \quad i = 1 \div nb. \quad (7)$$

- Xác định gradient của các khớp nối giữa lớp ẩn A và lớp ẩn B:

$$\text{grad}^{A_{ji}} = \delta^b_i \cdot O^{a_{ji}}, \quad j = 1 \div na, \quad i = 1 \div nb. \quad (8)$$

- Xác định gia số hiệu chỉnh:

$$\Delta w^{B_{ji}} = -\alpha \cdot \text{grad}^{A_{ji}} + \beta \cdot \Delta w^{B_{ji}}(l-1), \quad j = 1 \div na, \quad i = 1 \div nb, \quad (9)$$

trong đó, $\Delta w^{B_{ji}}(l-1)$ - là gia số hiệu chỉnh của vòng lặp trước.

- Hiệu chỉnh các véc tơ trọng số:

$$w^{B_{ji}} = w^{B_{ji}} + \Delta w^{B_{ji}}, \quad j = 1 \div na, \quad i = 1 \div nb. \quad (10)$$

Bước 5: Hiệu chỉnh ma trận trọng số giữa lớp vào X và lớp ẩn A:

- Xác định véc tơ delta của lớp ẩn A: $\delta^A = (\delta^a_1, \delta^a_2, \dots, \delta^a_{na})$,

$$\delta^a_i = (1 - O^a_i) \cdot O^a_i \cdot \sum_{j=1}^{nb} (w^{B_{ij}} \cdot \delta^b_j), \quad i = 1 \div na. \quad (11)$$

- Xác định gradient của các khớp nối giữa lớp vào X và lớp ẩn A:

$$\text{grad}^{X_{ji}} = \delta^a_i \cdot x_{ji}, \quad j = 1 \div n, \quad i = 1 \div na. \quad (12)$$

- Xác định gia số hiệu chỉnh:

$$\Delta w^{A_{ji}} = -\alpha \cdot \text{grad}^{X_{ji}} + \beta \cdot \Delta w^{A_{ji}}(l-1), \quad j = 1 \div n, \quad i = 1 \div na, \quad (13)$$

trong đó, $\Delta w^{A_{ji}}(l-1)$ - là gia số hiệu chỉnh của vòng lặp trước.

- Hiệu chỉnh các véc tơ trọng số:

$$w^{A_{ji}} = w^{A_{ji}} + \Delta w^{A_{ji}}, \quad j = 1 \div n, \quad i = 1 \div na. \quad (14)$$

Quay lại *bước 2*.

Bước 6: Kết thúc huấn luyện mạng.

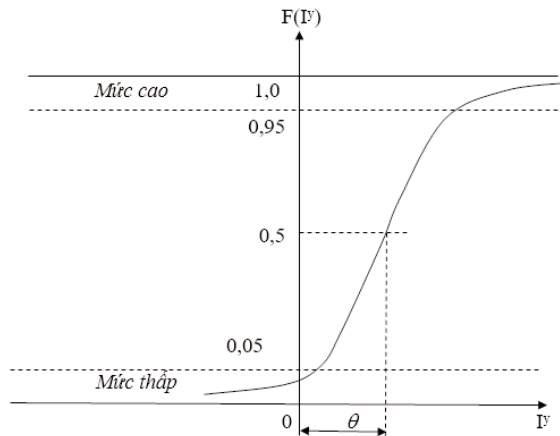
Như vậy, đối với một cặp mẫu, sau khi huấn luyện mạng ta thu được tập hợp các ma trận trọng số $W_s = (W^A, W^B, W^V)$. Thường số cặp mẫu bằng số nơ ron của lớp ra trong mạng. Huấn luyện cho tất cả các cặp mẫu, ta thu được các tập hợp ma trận trọng số của mạng nơ ron nhiều lớp.

4. Phân loại ảnh viễn thám bằng mạng nơ ron nhiều lớp

Giả sử cần phân loại m lớp các đối tượng theo ảnh viễn thám N kênh phổ (hoặc/và các dấu hiệu dẫn xuất). Dựa vào tính chất của các lớp đối tượng cần phân loại, xác định n kênh phổ (hoặc/và các dấu hiệu dẫn xuất) tương thích để sử dụng.

Mỗi lớp đối tượng được cho một bộ mẫu huấn luyện. Dựa vào các bộ mẫu, tiến hành huấn luyện mạng theo thuật toán trình bày ở trên. Kết quả ta thu được các tập hợp ma trận trọng số W_s .

Giả sử mạng nơ ron sử dụng hàm kích hoạt sigmoid, đầu ra có giá trị tiệm cận dưới và tiệm cận trên bằng 0 và 1 khi giá trị đầu vào của nơ ron bằng âm vô cùng ($-\infty$) và dương vô cùng ($+\infty$). Do vậy, cần phải đưa ra giá trị ngưỡng mức thấp và ngưỡng mức cao cho các giá trị của các nơ ron lớp ra gần với 0 và 1, ví dụ như, lấy ngưỡng dưới bằng 0,05 và ngưỡng trên bằng 0,95 (Hình 2).



Hình 2: Hàm kích hoạt sigmoid với độ chuyển dịch θ
 D – véc tơ đầu vào của lớp ra; $F(D)$ - véc tơ đầu ra.

Đối với mạng nơ ron nhiều lớp, hàm kích hoạt sigmoid dịch chuyển đi một giá trị $\theta = w_{n+1}$, là trọng số của khớp nối từ nơ ron chuyên vị của lớp ẩn cuối cùng đến các nơ ron của lớp ra (hằng số cho tất cả các nơ ron lớp ra).

Ký hiệu R_ω - là tập hợp các véc tơ có m phần tử, trong đó có duy nhất một phần tử có giá trị ở mức cao và các phần tử còn lại đều có giá trị ở mức thấp.

Mỗi pixel hoặc mỗi vùng ảnh (nếu ảnh đã được phân tích cụm) cho một véc tơ mở rộng đại diện đầu vào (Hình 1): $X = (x_1, x_2, \dots, x_n, 1)^T$, các phần tử x_i ($i = 1 \div n$) có thể là phản xạ phổ ở kênh i hoặc/và các dấu hiệu dẫn xuất.

Dựa vào các tập hợp ma trận trọng số W_s thu được sau khi huấn luyện mạng, xác định các véc tơ đầu ra theo véc tơ đầu vào X , ta thu được m véc tơ đầu ra Y_i ($i = 1 \div m$). Ký hiệu các lớp phân loại đối tượng là ω_i ($i = 1 \div m$).

Quy tắc phân loại như sau:

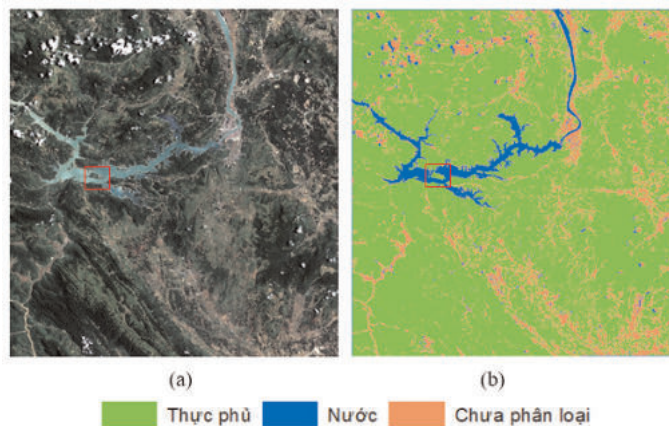
$$X \in \omega_i \ (i = 1 \div m) \iff Y_i \in R_\omega \text{ và } Y_k \notin R_\omega \ (k = 1 \div m, k \neq i). \quad (15)$$

Có thể mô tả như sau: Véc tơ đầu vào X thuộc lớp đối tượng ω_i ($i = 1 \div m$) khi và chỉ khi tồn tại duy nhất một véc tơ đầu ra Y_i , xác định theo bộ mẫu huấn luyện, thuộc tập hợp các véc tơ R_ω .

Nếu đầu ra của véc tơ X không tồn tại véc tơ $Y_i \in R_\omega$ ($i = 1 \div m$), thì pixel X không được phân loại. Nếu tồn tại lớn hơn hoặc bằng 2 véc tơ $Y_i \in R_\omega$ ($i = 1 \div m$), phân loại không xác định. Trường hợp này cần đánh giá lại mức độ phân biệt các mẫu huấn luyện. Trường hợp khác, pixel X nằm trên mặt phân chia các lớp đối tượng trong không gian dấu hiệu.

Thuật toán huấn luyện mạng nơ ron nhiều lớp và phương pháp phân loại ảnh viễn thám trên đây đã được các tác giả lập trình modul phần mềm bằng ngôn ngữ Matlab.

Kết quả phân loại đối tượng mặt nước và thực phủ bằng tư liệu ảnh viễn thám Landsat-8, số hiệu LC81270462020180LGN00 độ phân giải 30m, thu nhận ngày 28-6-2020 khu vực tỉnh Hòa Bình (Hình 3), đạt độ chính xác $k = 0,9$.



Hình 3: Kết quả phân loại ảnh viễn thám sử dụng mạng nơ ron nhân tạo hai lớp ẩn
(a). Ảnh trước phân loại, (b). Ảnh sau phân loại

5. Kết luận

Kết quả thử nghiệm thuật toán huấn luyện mạng nơ ron nhiều lớp trên đây và ứng dụng để phân loại ảnh viễn thám, cho thấy tính khả thi của việc bán tự động hóa cập nhật cơ sở dữ liệu không gian địa lý, thay thế cho các phương pháp giải đoán ảnh truyền thống. Đối với việc phân loại các lớp đối tượng khả tách tuyến tính, có thể sử dụng mạng nơ ron một lớp đơn giản. Đối với trường hợp các lớp đối tượng không hoàn toàn thỏa mãn tính chất khả tách tuyến tính, nên sử dụng mạng nơ ron nhiều lớp trình bày trên đây, cho độ chính xác cao.○

Tài liệu tham khảo

- [1]. Jwan Al-doski, Shattri B. Mansor and Helmi Zulhaidi Mohd Shafri (2013), *Image Classification in Remote Sensing*, Journal of Environment and Earth Science Vol. 3, No.10.
- [2]. Л.Н. Чабан (2014), *Автоматизированная обработка аэрокосмической информации для картографирования геопространственных данных*, Москва.
- [3]. Nguyễn Du Khang (2020), *Tài liệu hướng dẫn “Xử lý và giải đoán ảnh số viễn thám”*, Đề tài mã số TNMT.2018.08.08, Hà Nội.○

Summary

Detailed algorithm for remote sensing images classification by multi-layered artificial neural networks

Nguyen Du Khang, Nguyen Thi Thanh Binh

National Remote Sensing Department, MONRE

With the rapid development of remote sensing data acquisition and storage methods, today, a big data need to be processed and analyzed. These data are so large that the capabilities of the experts are no longer sufficient, leading to the continual increase in demand for automated data analysis methods. Application of artificial neural networks is one of the solutions to solve the above problem. The paper proposes a detailed algorithm for training multi-layered neural networks and remote sensing image classification application. Experimental results and evaluation of accuracy show that this method is quite feasible and simple to semi-automate the classification of objects represented on remote sensing images, for the establishment and updating of data national geospatial data.○