

# THUẬT TOÁN PHÂN LOẠI ẢNH VIỄN THÁM BẰNG MẠNG NƠI RON NHÂN TẠO MỘT LỚP

Nguyễn Dư Khang, Nguyễn Thị Thanh Bình, Nguyễn Thị Thu Trang  
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 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 một 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.

**Từ khóa:** Mạng nơ ron nhân tạo; Moment; Phân loại ảnh viễn thám.

## Abstract

### *Detailed algorithm for remote sensing image classification using single-layer neural networks*

*In context of the rapid development of remote sensing data acquisition and storage methods recently, big data should be processed and analyzed. Conventional image classification methods may not suit in analyzing such big data and thus there is an increasing demand for automated data analysis methods. Application of artificial neural networks is one of the solutions to solve the above issue. This paper proposes a detailed algorithm for training single-layer neural networks and its application for remote sensing image classification. Experimental results and accuracy assessment showed that this method was quite feasible and simple to semi-automate the classification of objects represented on remote sensing images, contributing in establishing and updating national geospatial data.*

**Keywords:** Neural network; Moment; Classification of remote sensing images.

## 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 tài liệu trong nước chủ yếu công bố kết quả ứng dụng mạng nơ ron trong phân loại ảnh viễn thám thông qua các thanh công cụ (Tool) có sẵn trong các phần mềm thương mại, việc đưa ra thuật toán chi tiết về vấn đề này chưa đề cập. Các công trình liên quan đến ứng dụng công nghệ

viễn thám để thành lập, cập nhật dữ liệu nền địa lý theo hướng bán tự động hóa hầu như thiếu vắng, phương pháp truyền thống hiện nay chủ yếu giải đoán ảnh viễn thám bằng mắt.

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 cấp Bộ Tài nguyên và Môi trường, mã số TNMT.2018.08.08). Bài báo này trình bày thuật toán chi tiết huấn luyện mạng nơ ron nhân tạo một lớp và ứng dụng phân

**Nghiên cứu**

loại ảnh viễn thám phục vụ cho công tác thành lập, cập nhật dữ liệu không gian địa lý quốc gia.

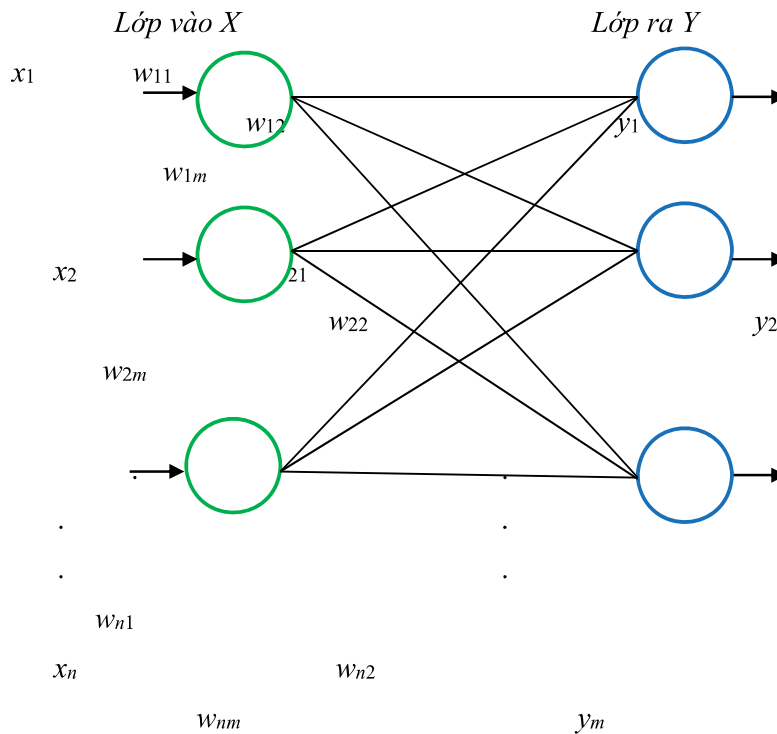
**2. Huấn luyện mạng**

Mạng nơ ron một lớp là dạng đơn giản nhất của các mạng nơ ron nhân tạo, nó chỉ gồm lớp vào và lớp ra, không có các lớp ẩn. Có nhiều mô hình mạng nơ ron một lớp, sau đây trình bày mạng Perceptron. Đó là mạng truyền thẳng có nhiều đầu ra, thường được sử dụng khi lớp ra bảo đảm

khả tách tuyến tính. Trường hợp ngược lại, sử dụng mạng nơ ron nhiều lớp.

Ký hiệu véc tơ đầu vào là  $X = (x_1, x_2, \dots, x_n)^T$  và véc tơ đầu ra là  $Y = (y_1, y_2, \dots, y_m)^T$ , trong đó,  $n$  - là số tín hiệu đầu vào, tương ứng với  $n$  nơ ron của lớp vào;  $m$  - là số nơ ron của lớp ra, tương ứng với  $m$  nơ ron của lớp ra (Hình 1).

Ký hiệu  $w_{ij}$  - là trọng số của khớp nối giữa nơ ron  $i$  của lớp vào và nơ ron  $j$  của lớp ra,  $i = 1 \div n; j = 1 \div m$ .



**Hình 1: Mạng nơ ron một lớp**

Ký hiệu  $W_i = (w_{1i}, w_{2i}, \dots, w_{ni})$ ,  $i = 1 \div m$  - là véc tơ trọng số của các khớp nối giữa các nơ ron lớp vào và lớp ra  $i$ .

Tập hợp các véc tơ trọng số trên tạo thành một ma trận trọng số:

$W = (W_1, W_2, \dots, W_m)^T$ , hay:

$$W = \begin{pmatrix} w_{11} & w_{21} & \dots & w_{n1} \\ w_{12} & w_{22} & \dots & w_{n2} \\ \dots & \dots & \dots & \dots \\ w_{1m} & w_{2m} & \dots & w_{nm} \end{pmatrix} \quad (1)$$

Khi đó, véc tơ đầu vào của lớp ra tính theo công thức:

$$y_i = F(I_i), i = 1 \div m, \quad (3)$$

Các phần tử của véc tơ đầu ra tính theo công thức:

Trong đó,  $F$  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:  $(e^{2x} - 1)/(e^{2x} + 1)$ .

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

Huấn luyện mạng là xác định ma trận trọng số  $W$ , 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 (cho một cặp mẫu) tiến hành như sau:

Bước 1: Nhập dữ liệu ban đầu

- Số nơ ron của lớp vào  $X_s$  và của lớp ra  $Y_s$ , tương ứng là  $n$  và  $m$ ;

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

- Tốc độ huấn luyện  $\alpha$ ; moment  $\beta$ ; giới hạn sai số  $\theta$  và giới hạn vòng lặp *Epoch*;

- Ma trận trọng số  $W$  lấy ngẫu nhiên trong khoảng  $[-0,5; 0,5]$ . Lấy  $l = 0$ .

Bước 2: Tính  $l = l + 1$ , nếu  $l > Epoch$ , chuyển đến *bước 3*.

- Xác định véc tơ đầu vào của lớp ra:

$$I = W \cdot X_s = (I_1, I_2, \dots, I_m)^T$$

- Xác định véc tơ đầu ra  $Y_s' = (y'_1, y'_2, \dots, y'_m)^T$ :

$$y'_i = F(I_i), i = 1 \div m,$$

$F$  là hàm kích hoạt.

Tính sai số bình phương trung bình theo công thức:

$$MSE = \frac{\sum_{i=1}^m (y_i - y'_i)^2}{m}, \quad (4)$$

nếu  $MSE \leq \theta$ , chuyển đến *bước 3*.

- Hiệu chỉnh ma trận trọng số  $W = (W_1, W_2, \dots, W_m)^T$ :

+ Xác định véc tơ gia số hiệu chỉnh trọng số, theo công thức:

$$\begin{aligned} \Delta W_i &= -\alpha \cdot X_s \cdot (y'_i - y_i) + \beta. \\ \Delta W_i(l-1), i &= 1 \div m, \end{aligned} \quad (5)$$

Trong đó,  $\Delta W_i(l-1)$  - là véc tơ gia số hiệu chỉnh của vòng lặp trước.

+ Xác định các véc tơ trọng số:

$$W_i = W_i + \Delta W_i, i = 1 \div m. \quad (6)$$

Quay lại *bước 2*.

Bước 3: Kết thúc huấn luyện mạng.

Như vậy, đối với một cặp mẫu, sau khi huấn luyện mạng ta thu được ma trận trọng số  $W$ . Thường số cặp mẫu bằng số nơ ron của lớp ra trong mạng.

Huấn luyện cho tất cả các cặp mẫu, ta thu được tập hợp các ma trận trọng số của mạng nơ ron.

### 3. Phân loại ảnh viễn thám bằng mạng nơ ron một lớp

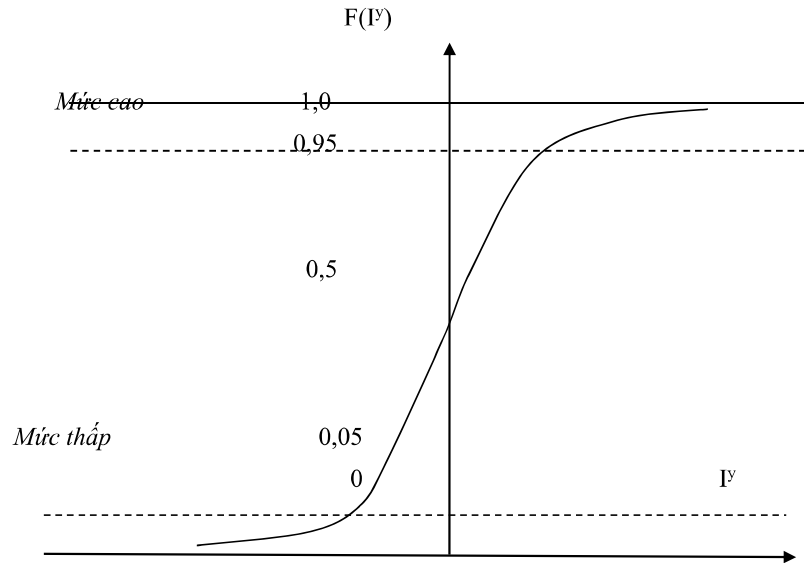
Mạng nơ ron một lớp được sử dụng trong phân loại ảnh viễn thám khi các lớp đối tượng phân loại thỏa mãn tính chất khả tách tuyến tính, ví dụ như: mặt nước, đất trống, thảm thực vật,...

Giả sử cần phân loại  $m$  lớp 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 tập hợp các ma trận trọng số  $W$ .

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).

## Nghiên cứu



**Hình 2: Hàm kích hoạt sigmoid mạng nơ ron một lớp**

$I^v$  - véc tơ đầu vào của lớp ra;  $F(I^v)$  - véc tơ đầu ra

Ký hiệu  $R_\omega$  - 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 tín hiệu đầu vào (Hình 1):  $X = (x_1, x_2, \dots, x_n)^T$ , các phần tử  $x_i$  ( $i = 1 \div n$ ) có thể là phân xạ phổ ở kênh  $i$  hoặc/và các dấu hiệu dẫn xuất của nó.

Dựa vào tập hợp các ma trận trọng số  $W$  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  $\omega_i$  ( $i = 1 \div m$ ).

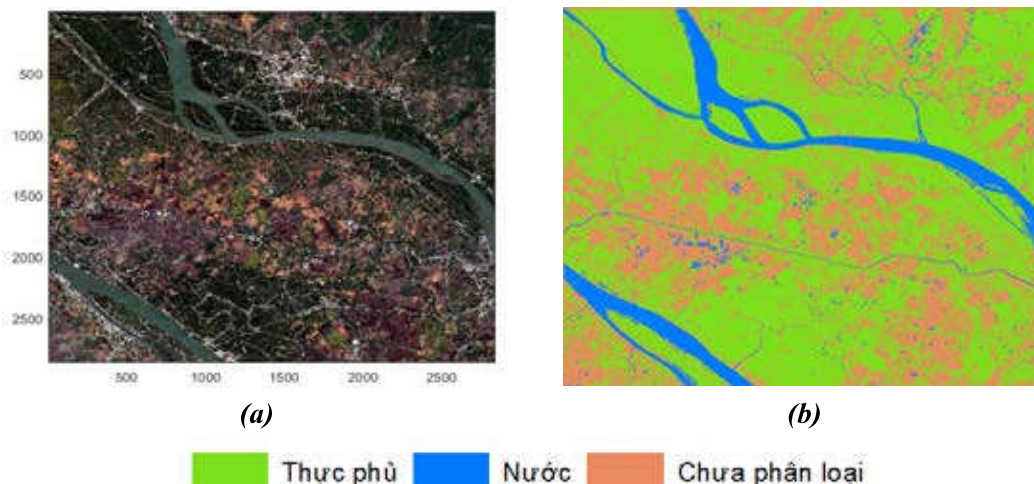
Quy tắc phân loại như sau:

$$\begin{aligned} X &\in \omega_i \quad (i = 1 \div m) \\ Y_i &\in R_\omega \text{ và } Y_k \in R_\omega \quad (k = 1 \div m, k \neq i). \end{aligned} \quad (7)$$

Có thể mô tả như sau: Véc tơ đầu vào  $X$  thuộc lớp đối tượng  $\omega_i$  ( $i = 1 \div m$ ) 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  $R_\omega$ .

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 một 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. So với các phương pháp phân loại ảnh viễn thám khác, phương pháp phân loại ảnh viễn thám cho các đối tượng mặt nước và thực vật là tối ưu nhất. Kết quả phân loại đối tượng mặt nước và thực vật bằng tư liệu ảnh viễn thám Sentinel-2A, số hiệu S2A\_MSIL2A\_N0214\_R118\_T48PWS độ phân giải 10 m, thu nhận ngày 25/02/2020 khu vực thành phố Cao Lãnh tỉnh Đồng Tháp (Hình 3), với các dấu hiệu giải đoán là giá trị phản xạ phổ của các pixel ảnh ở các kênh Blue, Green, Red và NIR, đạt độ chính xác  $\kappa = 0,92$ .



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

#### 4. Kết luận

Kết quả thử nghiệm thuật toán huấn luyện mạng nơ ron một 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 bằng mắt truyền thống. Đối với việc phân loại các lớp đối tượng khả tách tuyến tính, sử dụng mạng nơ ron một lớp đơn giản và cho độ chính xác cao. Đối với trường hợp các lớp đối tượng không hoàn toàn thỏa mãn tính chất khả tách tuyến tính, có thể sử dụng mạng nơ ron nhiều lớp hoặc các phương pháp phân loại ảnh phù hợp khác, nhóm tác giả sẽ lần lượt giới thiệu cùng

bạn đọc thuật toán chi tiết của các phương pháp này trong các nghiên cứu tiếp theo.

#### 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*. Московский государственный университет геодезии и картографии, Москва.
- BBT nhận bài: 15/10/2020; Phản biện xong: 02/11/2020; Chấp nhận đăng: 15/12/2020