



Original Article

# Application of Deep Learning Algorithm to Build an Automated Cloud Segmentation Model Based on Open Data Cube Framework

Pham Vu Dong<sup>1</sup>, Bui Quang Thanh<sup>1</sup>, Nguyen Quoc Huy<sup>1</sup>, Vo Hong Anh<sup>2</sup>,  
Pham Van Manh<sup>1,\*</sup>

<sup>1</sup>*VNU University of Science, 334 Nguyen Trai, Hanoi, Vietnam*

<sup>2</sup>*Central Remote Sensing Station, National Remote Sensing Department,  
79 Van Tien Dung, Tu Liem, Hanoi, Vietnam*

Received 11 September 2019

Revised 23 April 2020; Accepted 28 August 2020

**Abstract:** Cloud detection is a significant task in optical remote sensing to reconstruct the contaminated cloud area from multi-temporal satellite images. Besides, the rapid development of machine learning techniques, especially deep learning algorithms, can detect clouds over a large area in optical remote sensing data. In this study, the method based on the proposed deep-learning method called ODC-Cloud, which was built on convolutional blocks and integrating with the Open Data Cube (ODC) platform. The results showed that our proposed model achieved an overall 90% accuracy in detecting cloud in Landsat 8 OLI imagery and successfully integrated with the ODC to perform multi-scale and multi-temporal analysis. This is a pioneer study in techniques of storing and analyzing big optical remote sensing data.

**Keywords:** Optical Remote Sensing, Landsat 8 OLI, automatic cloud detection, deep-learning, Open Data Cube.

\* Corresponding author.

*E-mail address:* [manh10101984@gmail.com](mailto:manh10101984@gmail.com)

<https://doi.org/10.25073/2588-1094/vnuees.4441>

# Ứng dụng thuật toán học máy sâu xây dựng mô hình tự động phát hiện vùng mây trên nền tảng dữ liệu khối

Phạm Vũ Đông<sup>1</sup>, Bùi Quang Thành<sup>1</sup>, Nguyễn Quốc Huy<sup>1</sup>, Võ Hồng Anh<sup>2</sup>,  
Phạm Văn Mạnh<sup>1,\*</sup>

<sup>1</sup>Trường Đại học Khoa học Tự nhiên, ĐHQGHN, 334 Nguyễn Trãi, Hà Nội, Việt Nam

<sup>2</sup>Đài Viễn thám Trung ương, Cục Viễn thám Quốc gia, 79 Văn Tiến Dũng, Bắc Từ Liêm, Hà Nội, Việt Nam

Nhận ngày 11 tháng 09 năm 2019

Chỉnh sửa ngày 23 tháng 4 năm 2019; Chấp nhận đăng ngày 28 tháng 8 năm 2020

**Tóm tắt:** Loại bỏ vùng mây trên ảnh viễn thám quang học là một trong những bước đầu tiên trong quá trình tiền xử lý ảnh, nhằm phục vụ tái cấu trúc dữ liệu những vùng bị che phủ bởi mây từ dữ liệu ảnh vệ tinh đa thời gian. Để giải quyết vấn đề này, một mô hình được thử nghiệm với thuật toán dựa trên học máy sâu được đề xuất trong bài viết này. Mô hình thuật toán này được phát triển trên nguyên lý sử dụng một mạng tích chập (convolutional blocks) lên nền tảng dữ liệu khối (Open Data Cube) được đào tạo bởi nhiều phân mảnh của ảnh Landsat 8 OLI. Mô hình này, được gọi là ODC-Cloud có khả năng phát hiện vùng mây trên diện rộng và cục bộ trong một hình ảnh bằng cách sử dụng các khối chập. Vì mô hình được đề xuất là một giải pháp đầu cuối không yêu cầu bước tiền xử lý phức tạp. Kết quả thử nghiệm của nghiên cứu này đạt độ chính xác trên 90% trong việc phát hiện và loại bỏ vùng mây khỏi ảnh và tích hợp thành công mô hình lên nền tảng dữ liệu khối để thực hiện các phân tích đa tỉ lệ và đa thời gian. Đây là một trong những giải pháp mới trong việc lưu trữ và xử lý dữ liệu ảnh viễn thám quang học trên nền tảng dữ liệu lớn.

**Từ khóa:** Viễn thám quang học, Landsat 8 OLI, tự động phát hiện mây, học máy sâu, dữ liệu khối.

## 1. Mở đầu

Với sự phát triển nhanh của công nghệ viễn thám, ảnh viễn thám quang học được sử dụng rộng rãi trong nhiều lĩnh vực như so sánh đối chiếu thực địa, phân loại và theo dõi biến động lớp phủ/sử dụng đất [1]. Tuy nhiên phần lớn ảnh viễn thám quang học đều có sự xuất hiện của vùng mây với độ che phủ khác nhau. Những khu vực bị mây che phủ trên ảnh không chỉ gây ra sự mất thông tin mà còn gây ra khó khăn trong việc xác định các đối tượng bên dưới vùng mây [2], [3]. Ngoài ra, các ảnh viễn thám quang học với độ che phủ mây nhiều có thể khiến dữ liệu ảnh không có giá trị để phân tích bề mặt lớp phủ. Tuy

nhiên, các thông tin về giá trị của mây cũng có thể cung cấp các tham số khác phục vụ nghiên cứu thời tiết, thảm họa tự nhiên như giông bão, núi lửa phun trào [4]. Do đó, phát hiện vùng mây là một trong những vấn đề nghiên cứu quan trọng để quan sát Trái Đất bằng hình ảnh viễn thám quang học. Đây là một trong những bước quan trọng trong quá trình tiền xử lý ảnh, trước khi phục vụ cho các mục đích nghiên cứu và phân tích sau này.

Nhiều phương pháp và mô hình phát hiện vùng mây đã được nghiên cứu và thử nghiệm trong những thập kỷ gần đây. Hầu hết các phương pháp phát hiện được thiết kế cho các loại

\* Tác giả liên hệ.

E-mail address: manh10101984@gmail.com

<https://doi.org/10.25073/2588-1094/vnuees.4441>

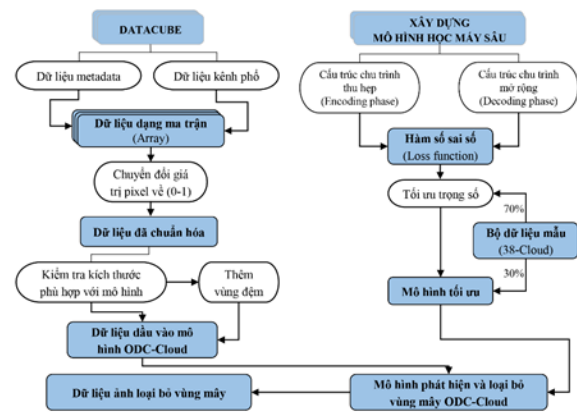
ảnh vệ tinh đa phổ. Với các dải hồng ngoại nhiệt hoặc sóng ngắn, rất dễ phân biệt các vùng mây và không mây. Các phương pháp này đã được áp dụng thành công cho các dữ liệu viễn thám khác nhau theo độ phân giải không gian và thời gian. Có thể được chia làm ba phương pháp chính: (1) Phương pháp phân ngưỡng [5,6], các tác giả đã sử dụng phương pháp F-mask (function of mask) để lọc những vùng có khả năng là mây ra khỏi những vùng không mây, ngoài ra phương pháp này còn sử dụng kênh ảnh Cirrus (kênh 9 của ảnh Landsat 8 OLI) để tăng độ chính xác việc xác định vùng mây, và hiện nay được sử dụng để tạo ra kênh lọc mây trên ảnh Landsat 8 OLI; (2) Phương pháp thủ công [7], tác giả sử dụng phương pháp HOT (haze optimized transformation) và sử dụng liên kết giữa hai kênh ảnh Landsat để phân biệt vùng mây mỏng và không mây; và (3) Phương pháp học máy sâu [8-10] được sử dụng trong một vài năm trở lại đây, các tác giả sử dụng mạng nơ-ron tích chập Convolutional Neural Networks (CNNs) cho việc phân biệt vùng mây dày, mỏng và không mây. Phương pháp học máy sâu sử dụng cấu trúc U-Net cho thấy sự vượt trội trong các ứng dụng phân tích ảnh [11-14].

Dựa trên những cơ sở này và nhu cầu sử dụng ảnh vệ tinh đa thời gian ngày càng lớn, có thể thấy được khả năng phát triển một thuật toán học máy sâu trong việc ứng dụng xây dựng mô hình tự động phát hiện và loại bỏ vùng mây trên ảnh viễn thám quang học. Đồng thời với sự phát triển nhanh của công nghệ viễn thám, việc lưu trữ và phân tích dữ liệu lớn đã trở thành một trong những bài toán không chỉ trong lĩnh vực viễn thám mà còn trong nhiều lĩnh vực khác [9]. Open Data Cube (ODC) là một giải pháp cho việc truy cập, quản lý và phân tích dữ liệu địa lý lớn nói chung và dữ liệu viễn thám nói riêng. Môi trường ODC cho phép liên kết các dữ liệu thành dạng lưới dữ liệu theo chiều không gian và dạng khối theo chiều thời gian. Chính vì đặc điểm này, việc phân tích xây dựng mô hình xử lý ảnh viễn thám trên nền tảng ODC có khả năng áp dụng trên nhiều khu vực, nhiều giai đoạn khác nhau và khả năng tính toán nhanh với bộ dữ liệu lớn. Xuất phát từ những nhu cầu và điều kiện trên, nghiên

cứu này bước đầu áp dụng và xây dựng thuật toán học máy sâu (deep-learning) để mô hình hóa tự động phát hiện và loại bỏ vùng mây trên ảnh viễn thám quang học (Landsat 8 OLI) trên nền tảng ODC với tên gọi là ODC-Cloud.

**2. Phương pháp nghiên cứu**

Trong phần này, nghiên cứu trình bày thuật toán đề xuất trong mô hình ODC-Cloud để phát hiện và loại bỏ vùng mây cho ảnh Landsat 8 OLI. Sơ đồ phương pháp của mô hình được thể hiện trong Hình 1.



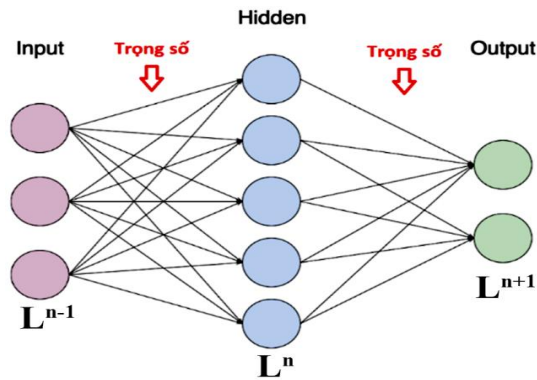
Hình 1. Sơ đồ mô hình thuật toán phát hiện và loại bỏ vùng mây ODC-Cloud.

**2.1. Phương pháp học máy sâu**

Phương pháp học máy sâu được phát triển từ mô hình mạng nơ-ron (Neural Network). Một trong những phương pháp học máy sâu tiêu biểu trong việc phân tích các đối tượng hình ảnh. Phương pháp mạng nơ-ron tích chập sâu Deep Convolutional Neural Network (DCNN) trong phân tích ảnh thể hiện khả năng tự động phân loại (automatic classification), phân đoạn ảnh (segmentation) và phát hiện đối tượng (object detection). Trong nghiên cứu này, mô hình thử nghiệm phát hiện mây có chức năng tự động xác định và loại bỏ vùng có mây bằng cách phân mảnh các đối tượng trên ảnh (phương trình 1).

$$Y = f(X; \theta) \tag{1}$$

Với dữ liệu ảnh  $X \in R^{w,h,c}$ , mô hình sẽ tái cấu trúc ảnh  $Y \in R^{w,h,1}$  với  $w$  là chiều rộng;  $h$  là số hàng của pixel; và  $c$  là số kênh ảnh. Khi đó, dữ liệu đầu ra  $Y$  có giá trị (0-1), đại diện cho sự xuất hiện của vùng mây trên ảnh, và  $\theta$  là các trọng số của mô hình, đối với mạng nơ-ron (Neural Network) các giá trị trọng số  $\theta$  của lớp  $L^{n-1}$  được sử dụng để xác định giá trị của các nốt (node) trong lớp  $L^n$  (Hình 2).



Hình 2. Mạng nơ-ron (Neural Network) với lớp đầu vào (Input), lớp ẩn (Hidden) và lớp đầu ra (Output).

Trong đó:  $Hidden = f_1(Input; \theta_1)$ ;  $Output = f_2(Hidden; \theta_2)$ . Bằng cách hợp hàm  $f$  từ hai hàm  $f_1$  và  $f_2$ , các trọng số được biểu diễn theo (phương trình 2).

$$Y = f_2(f_1(X; \theta_1); \theta_2) \quad (2)$$

Như vậy, trong một phương pháp học máy sâu (deep-learning) với mỗi dữ liệu đầu vào thì dữ liệu kết quả sẽ được tính toán qua rất nhiều hàm và mỗi kết quả của lớp này sẽ là dữ liệu đầu vào của lớp tiếp theo.

### 2.2. Cấu trúc mô hình ODC-Cloud

#### • Dữ liệu ảnh vệ tinh

Dữ liệu ảnh vệ tinh đa phổ Landsat 8 OLI bao gồm 09 kênh ảnh được thu nhận từ cảm biến OLI (Operational Land Imager) và 02 kênh ảnh từ cảm biến TIRS (Thermal Infrared Sensor). Trong nghiên cứu này, nhóm tác giả sử dụng 04 kênh ảnh phổ biến gồm Blue, Green, Red và Near-Infrared (Bảng 1). Với đặc điểm nhiều kênh ảnh hơn các loại ảnh truyền thống khác thường được sử dụng 03 kênh ảnh (Red, Green,

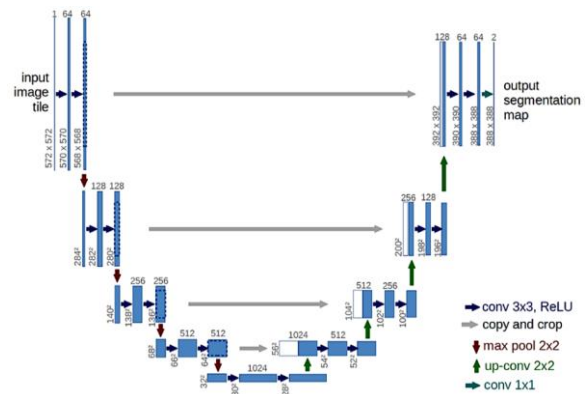
Blue) trong các ứng dụng phân tích ảnh sử dụng phương pháp học máy sâu (deep-learning) [15, 16]. Dữ liệu ảnh vệ tinh nói chung, hay ảnh Landsat 8 OLI nói riêng với đặc điểm có nhiều kênh ảnh, khả năng phát hiện đối tượng không chỉ dựa trên đặc điểm hình dạng mà còn dựa trên tính chất của đối tượng. Nghiên cứu này, ngoài 03 kênh ảnh (Red, Green, Blue) được sử dụng như các phương pháp học máy sâu khác. Mô hình ODC-Cloud sử dụng thêm kênh cận hồng ngoại (Near-Infrared) với mục đích cung cấp thêm thông tin về hình ảnh phản chiếu cũng như tính chất vật lý giúp việc phân đoạn ảnh được chính xác hơn.

Bảng 1. Các kênh phổ của ảnh Landsat 8 OLI được sử dụng trong nghiên cứu

Spectral Bands	Wavelength ( $\mu\text{m}$ )
Band 2 – Blue	0,452 – 0,512
Band 3 – Green	0,533 – 0,590
Band 4 – Red	0,636 – 0,673
Band 5 – Near-Infrared (Nir)	0,851 – 0,879

#### • Thiết lập mô hình ODC-Cloud

DCNN (Deep Convolutional Neural Network) được ứng dụng trong phân tích ảnh và dựa trên cấu trúc U-Net là một trong những cấu trúc tiêu biểu trong phân đoạn ảnh [11-14]. Mô hình cấu trúc U-Net được minh họa trong Hình 3.



Hình 3. Cấu trúc U-Net được sử dụng trong DCNN.

Với bộ lọc tích chập tại vị trí  $i, j$  của lớp đầu vào  $x$ , lớp đầu ra  $h_{i,j}$  được xác định bởi (phương trình 3).

$$h_{i,j} = \sum_{k=1}^m \sum_{l=1}^m w_{k,l} x_{i+k-1,j+l-1} \quad (3)$$

Trong đó  $m$  là kích thước chiều dài và chiều rộng của bộ lọc;  $h$  là lớp đầu ra;  $x$  là lớp đầu vào; và  $w$  là trọng số của bộ lọc.

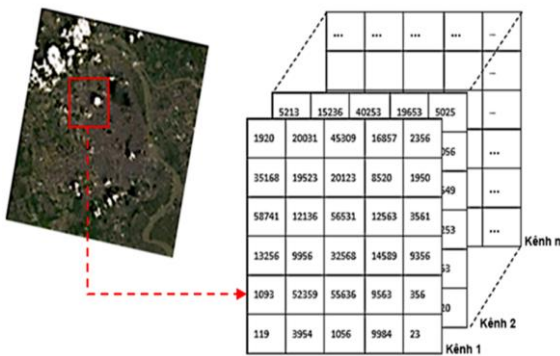
Số lượng bộ lọc sẽ quyết định số lượng kênh của lớp đầu ra và thường được chọn một cách thủ công. Bộ lọc chỉ được áp dụng cho một phần của toàn bộ lớp đầu vào nhưng toàn bộ kênh của đối tượng đầu vào đều được áp dụng. Bộ lọc biến đổi lớp đầu vào theo một hàm tuyến tính, theo sau là một sự phi tuyến (thường là một hàm phi tuyến). Hàm ReLU (rectified linear unit) được dùng để chuyển tất cả các giá trị âm thành giá trị 0 (phương trình 4).

$$ReLU(X) = \max(0, X) \quad (4)$$

Bước tiếp theo hàm sai số nhị phân chéo (binary cross-entropy) được áp dụng để tính sai số giữa kết quả mẫu và kết quả dự đoán vùng có mây (phương trình 5).

$$Loss(\hat{Y}, Y) = \frac{-1}{w,h} \sum_w \sum_h \hat{Y}_{w,h} \ln(Y_{w,h}) + (1 - \hat{Y}_{w,h}) \ln(1 - Y_{w,h}) \quad (5)$$

Với  $\hat{Y}$  là dữ liệu chia vùng mây mẫu và  $Y$  là kết quả dự đoán chia vùng mây. Hàm sai số này được sử dụng để tính toán giá trị độ dốc (gradient) của các giá trị trọng số trong  $\theta$ . Giá trị độ dốc sau đó sẽ được sử dụng để tối ưu hóa mô hình.

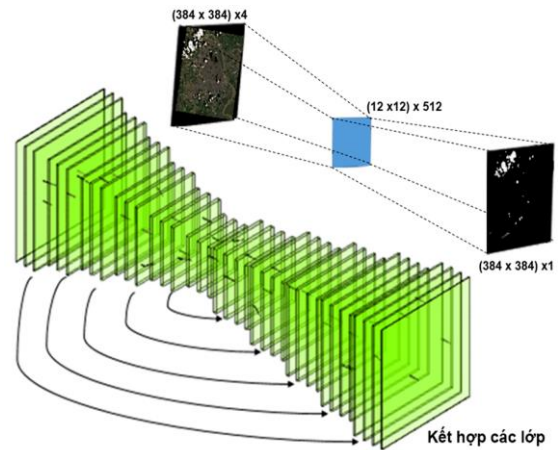


Hình 4. Các giá trị pixel trên ảnh vệ tinh Landsat 8 OLI có giá trị khác nhau trên từng kênh ảnh.

Hình 5 thể hiện cấu trúc học sâu DCNN được đề xuất trong mô hình ODC-Cloud. Dữ liệu đầu

vào (Input) là ảnh có kích thước  $(384 \times 384) \times 4$ . Trong đó, 4 là số kênh ảnh (Red, Green, Blue, Near-Infrared).

Dữ liệu đầu vào ban đầu sẽ qua 2 bộ lọc tích chập  $(3 \times 3)$  (convolution layers) theo sau một lớp tổng hợp lớn nhất (max pooling layers). Qua mỗi một chu trình như vậy, kích thích (số hàng và số cột) của ảnh sẽ giảm đi một nửa và số lượng kênh ảnh sẽ tăng lên.



Hình 5. Mô hình DCNN dựa trên cấu trúc U-Net.

Qua 6 chu trình, ảnh có kích thước  $(12 \times 12) \times 512$ , sau đó ảnh sẽ đi qua chu trình tích hợp ngược và kết hợp lớp. Bộ lọc tích hợp ngược có kích thước  $(3 \times 3)$ , qua mỗi phép tích chập kích thước ảnh tăng lên gấp đôi và số kênh giảm đi một nửa (Bảng 2). Ảnh sau đó sẽ được kết hợp với lớp tương ứng ở giai đoạn trước đó và tiếp tục qua 2 bộ lọc tích chập  $(3 \times 3)$ .

Tương tự qua các chu trình mở rộng, đối tượng ảnh sẽ mở rộng ra kích thước  $(384 \times 384)$  như kích thước gốc ban đầu tuy nhiên số kênh chỉ còn lại một kênh với giá trị pixel từ (0-1). Hàm sigmoid sẽ được áp dụng cho lớp cuối cùng để xác định khả năng có mây trên ảnh (phương trình 6).

$$A = \frac{1}{1 + e^{-x}} \quad (6)$$

Trong đó  $A$  là giá trị Sigmoid của lớp cuối cùng trong khoảng  $(0 - 1)$ ;  $x$  là giá trị của lớp cuối cùng.

Bảng 2. Cấu trúc phương pháp học máy sâu DCNN áp dụng trong mô hình ODC-Cloud.

<b>Chu trình thu hẹp (Encoding phase)</b>	
Chu trình 1	384x384x4 -> convolution (3x3) 16 feature maps -> 384x384x16 -> convolution (3x3) 16 feature maps -> 384x384x16 -> max pooling (2x2) -> 192x192x16
Chu trình 2	192x192x16 -> convolution (3x3) 32 feature maps -> 192x192x32-> convolution (3x3) 32 feature maps -> 192x192x32 -> max pooling (2x2) -> 96x96x32
Chu trình 3	96x96x32 -> convolution (3x3) 64 feature maps -> 96x96x64-> convolution (3x3) 64 feature maps -> 96x96x64 -> max pooling (2x2) -> 48x48x64
Chu trình 4	48x48x64-> convolution (3x3) 128 feature maps -> 48x48x128-> convolution (3x3) 128 feature maps -> 48x48x128 -> max pooling (2x2) -> 24x24x128
Chu trình 5	24x24x128-> convolution (3x3) 256 feature maps -> 24x24x256-> convolution (3x3) 256 feature maps -> 24x24x256 -> max pooling (2x2) -> 12x12x256
Chu trình 6	12x12x256-> convolution (3x3) 512 feature maps -> 12x12x512-> convolution (3x3) 512 feature maps -> 12x12x512
<b>Chu trình mở rộng (Decoding phase)</b>	
Chu trình 7	12x12x512-> up convolution (3x3) 256 feature maps -> 24x24x256-> kết hợp lớp tại quá trình 5 -> 24x24x512-> convolution (3x3) 256 feature maps -> 24x24x256-> convolution (3x3) 256 feature maps -> 24x24x256

Chu trình 8	24x24x256-> up convolution (3x3) 128 feature maps -> 48x48x128-> kết hợp lớp tại quá trình 4 -> 48x48x256-> convolution (3x3) 128 feature maps -> 48x48x128-> convolution (3x3) 128 feature maps -> 48x48x128
Chu trình 9	48x48x128-> up convolution (3x3) 64 feature maps -> 96x96x64-> kết hợp lớp tại quá trình 3 -> 96x96x128-> convolution (3x3) 64 feature maps -> 96x96x64-> convolution (3x3) 64 feature maps -> 96x96x64
Chu trình 10	192x192x64-> up convolution (3x3) 32 feature maps -> 192x192x32-> kết hợp lớp tại quá trình 2 -> 192x192x64-> convolution (3x3) 32 feature maps -> 192x192x32-> convolution (3x3) 32 feature maps -> 192x192x32
Chu trình 11	192x192x32-> up convolution (3x3) 16 feature maps -> 384x384x16-> kết hợp lớp tại quá trình 1 -> 384x384x32-> convolution (3x3) 16 feature maps -> 384x384x16-> convolution (3x3) 1 feature maps (Sigmoid) -> 384x384x1

2.3. Dữ liệu mẫu và quá trình học máy

Mô hình ODC-Cloud sử dụng bộ dữ liệu mới cho mục đích phát hiện vùng mây trên ảnh vệ tinh Landsat 8 OLI. Bộ dữ liệu mẫu sử dụng là bộ dữ liệu mẫu 38-Cloud được lọc mây một cách thủ công [15]. Mô hình ODC-Cloud sử dụng 8.192 mẫu và được chia làm hai phần: (i) dữ liệu học máy (khoảng 70% số mẫu) và (ii) dữ liệu kiểm chứng (khoảng 30% số mẫu). Mỗi mẫu ảnh được phân thành các mảnh có kích thước (384 × 384). Các mảnh sẽ được sắp xếp theo thứ tự ngẫu nhiên để tối ưu hóa quá trình học máy. Giá trị pixel của dữ liệu đầu vào được tiêu chuẩn hóa về khoảng giá trị (0-1) (phương trình 7).

$$A = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{7}$$



Trong đó  $X$  là giá trị pixel;  $X_{min}$  là giá trị pixel nhỏ nhất trong ảnh,  $X_{max}$  là giá trị pixel lớn nhất trong ảnh.

Khả năng học máy (Learning rate) khởi điểm với giá trị là 0,0001 và kỹ thuật tối ưu hóa để giảm tỷ lệ học tập (Learning rate decay) được áp dụng, khả năng học máy sẽ giảm 10% qua mỗi lần và toàn bộ dữ liệu học máy được đi qua mô hình. Giá trị độ dốc (Gradient) sẽ được tính cho tám bộ mảnh dữ liệu được đi qua mô hình. Do đó, với mỗi bộ thì mô hình sẽ có 128 lần học máy và tối ưu trọng số. Quá trình học máy trong mô hình ODC-Cloud, được học toàn bộ dữ liệu là 20 lần, số lần tối ưu trọng số  $(20 \times 8) \times 128 = 20.480$  lần. Mô hình ODC-Cloud được tối ưu trên máy trạm CPU Xeon 2650V2 với card đồ họa Geforce GTX 1050Ti với khoảng thời gian xử lý 4 giờ và mô hình toán học được chạy trên nền tảng mã nguồn mở Python 3.6, kết hợp với deep learning framework Tensorflow 1.13 trong quá trình tối ưu mô hình. Các quá trình đọc ảnh và xuất dữ liệu ảnh vệ tinh được định dạng (\*.TIFF) được thực hiện qua thư viện GDAL Python.

#### 2.4. Đánh giá độ chính xác của mô hình ODC-Cloud

Sau quá trình học máy, hiệu suất của mô hình được đánh giá độ chính xác tổng thể (Overall Accuracy - OA) (phương trình 8). Hai ảnh gốc Landsat 8 OLI và trên dữ liệu khối tại Việt Nam (Vietnam Data Cube). Kết quả lớp đầu ra là ảnh có cùng kích thước với ảnh đầu vào và số lượng kênh ảnh là một, giá trị pixel nằm trong giá trị (0, 1). Hàm sigmoid đưa giá trị cuối cùng về giá trị 0 và giá trị 1, qua quá trình tối ưu mô hình, các dữ liệu từ vùng mây sẽ có giá trị tiến về 1 và các vùng không mây sẽ có giá trị tiến về 0. Kết quả dự đoán pixel vùng mây có xác suất  $P$  và pixel không mây là  $1 - P$ . Giá trị pixel sẽ được xác định là 1 ( $P > 0,5$ ) và là 0 nếu ( $P < 0,5$ ).

$$OA = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

Ở đây  $TP$ ,  $TN$ ,  $FP$  và  $FN$  là các chỉ số của ma trận được mô tả trong ma trận sai lầm dựa trên tính toán chỉ số Jaccard. Đây là số liệu được chấp nhận rộng rãi để đo hiệu suất của nhiều thuật toán phân đoạn ảnh [17] (Hình 6).

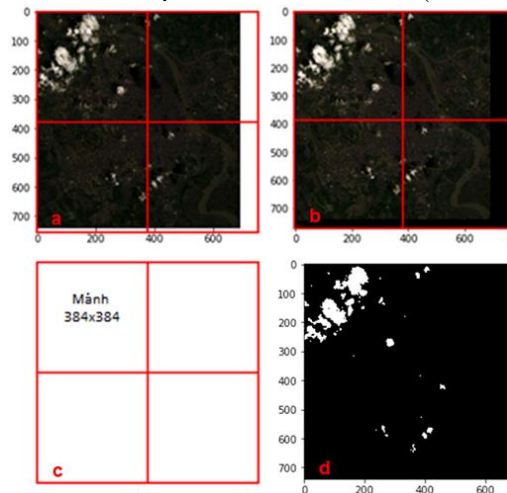
		Predicted class	
		P	N
Actual Class	P	True Positives (TP)	False Negatives (FN)
	N	False Positives (FP)	True Negatives (TN)

Hình 6. Ma trận sai lầm.

### 3. Kết quả thử nghiệm

#### 3.1. Mô hình ODC-Cloud trên dữ liệu khối tại Việt Nam

Xử lý toàn bộ vùng dữ liệu ảnh trên dữ liệu khối qua phương pháp DCNN (Deep Convolutional Neural Network) gặp nhiều khó khăn vì kích thước của các vùng khó lựa chọn đúng với kích thước dữ liệu đầu vào của mô hình. Trong nghiên cứu này, nhóm tác giả thêm một vùng đệm cho ảnh trước khi chia ảnh ra thành từng mảnh nhỏ để xử lý. Hình 7a thể hiện khu vực trên dữ liệu khối có kích thước  $(739 \times 691)$ , được thêm lớp vùng đệm cho cả số hàng và cột (các vùng đệm có giá trị pixel = 0) thành  $(768 \times 768)$  (Hình 7b). Như vậy khu vực được chia là bốn mảnh có kích thước nhỏ hơn ban đầu  $(384 \times 384)$  để mô hình tính toán và xử lý (Hình 7c). Sau khi xác định vùng mây được mô hình tính toán, phần vùng đệm này sẽ được loại bỏ và kết quả ảnh sẽ trở lại kích thước ban đầu (Hình 7d).



Hình 7. Khu vực được lựa chọn trên dữ liệu khối (7a), thêm vùng đệm trên ảnh (7b), chia mảnh (7c) và kết quả tách vùng mây (7d).

### 3.2. Độ chính xác và kết quả của mô hình

Với bộ dữ liệu huấn luyện, độ chính xác của mô hình được tính cho tám bộ mảnh, mỗi bộ chứa 1.024 ảnh với 04 kênh ảnh (Red, Green, Blue, Near-Infrared), độ chính xác tổng thể trung bình của tám bộ mảnh được thể hiện (Bảng 3).

Bảng 3. Độ chính xác của mô hình phát hiện vùng mây

Bộ mảnh	Độ chính xác (%)
1	96,96
2	97,16
3	97,31
4	97,30
5	97,22
6	97,14
7	97,45
8	97,22
Độ chính xác tổng thể (OA)	97,22

Hình 7 (bên phải) thể hiện ảnh vùng mây được tách khỏi ảnh Landsat 8 OLI từ dữ liệu khối tại Việt Nam (Vietnam Data Cube) thử nghiệm trên mô hình ODC-Cloud với độ chính xác tổng thể đạt được 90,26% (Bảng 4).

Bảng 4. Độ chính xác của mô hình ODC-Cloud loại bỏ vùng mây trên hai ảnh Landsat 8 OLI

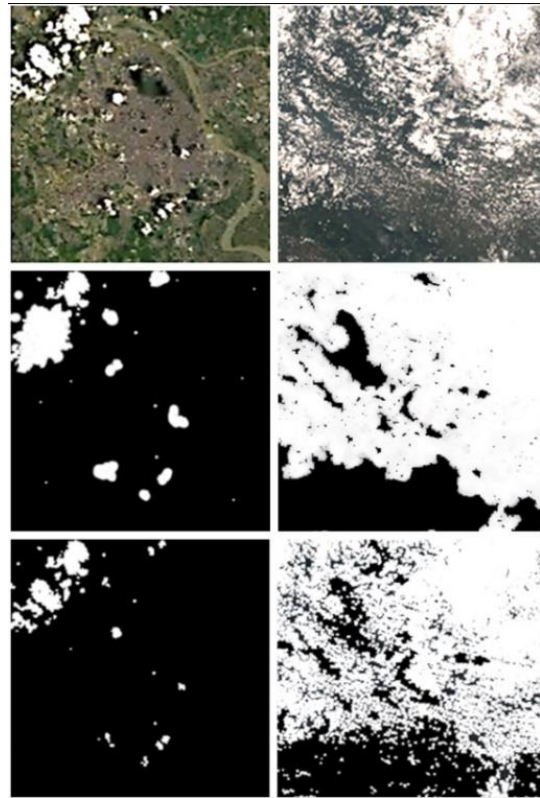
Ảnh	Độ chính xác
1	91,13 %
2	89,39 %
Độ chính xác tổng thể (OA)	90,26 %

Bảng 4 và Hình 8 cho thấy kết quả thử nghiệm mô hình ODC-Cloud trên dữ liệu khối tại Việt Nam với độ chính xác của từng ảnh thử nghiệm đều xấp xỉ đạt 90% và độ chính xác tổng thể cũng đạt 90%, điều đó cho thấy bước đầu mô hình ODC-Cloud có thể chiết tách các vùng mây trên ảnh vệ tinh Landsat 8 OLI với bộ dữ liệu lớn và có khả năng áp dụng chiết tách các vùng mây dày và trung bình. Hạn chế của mô hình thử nghiệm phát hiện và chiết tách vùng mây (ODC-Cloud) là chưa tính toán thêm vùng bóng mây và mây mờ trên ảnh. Trong nghiên cứu tiếp theo, nhóm tác giả sẽ thử nghiệm mô hình ODC-Cloud trên các loại ảnh viễn thám quang học khác và

tích hợp thuật toán phát hiện vùng mây, mây mỏng và bóng mây với bộ dữ liệu lớn hơn.

### 4. Kết luận

Trong nghiên cứu này, phương pháp được thử nghiệm dựa trên tiếp cận học máy sâu (deep learning) được áp dụng cho mô hình ODC-Cloud phát hiện các pixel vùng mây trên ảnh vệ tinh Landsat 8 OLI chỉ sử dụng 4 kênh ảnh (Red, Green, Blue, Near-Infrared). Kết quả đạt độ chính xác trên 90% phát hiện và chiết tách vùng mây khỏi ảnh vệ tinh Landsat 8 OLI. Mô hình ODC-Cloud có thể hỗ trợ đắc lực cho công tác tiền xử lý ảnh viễn thám quang học với bộ dữ liệu khối lớn tại Việt Nam và khả năng tính toán theo phân mảnh độc lập, mô hình thử nghiệm trong nghiên cứu này có thể áp dụng kỹ thuật tính toán song song cho một khu vực lớn và rất lớn.



Hình 8. Ảnh tổ hợp màu tự nhiên của hai ảnh Landsat 8 OLI (hàng đầu), ảnh vùng mây kiểm chứng (giữa), và ảnh kết quả vùng mây được chiết tách từ mô hình ODC-Cloud (bên dưới).



## Lời cảm ơn

Nghiên cứu này được hỗ trợ bởi Trung tâm Vũ trụ Việt Nam (VNSC) và tập đoàn I.M. System Group, Inc (<https://img.com>) trong việc cung cấp môi trường và bộ dữ liệu khối tại Việt Nam (Vietnam Data Cube) trong quá trình nghiên cứu và thử nghiệm. Các tác giả xin chân thành cảm ơn người đánh giá ẩn danh cho các ý kiến mang tính xây dựng và sâu sắc đã cải thiện nghiên cứu này.

## Tài liệu tham khảo

- [1] M. Shi, F. Xie, Y. Zi, J. Yin, Cloud detection of remote sensing images by deep learning, IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, 2016 (701-704) <https://doi.org/10.1109/IGARSS.2016.7729176>.
- [2] X. Jin, J. Li, T.J. Schmit, J. Li, M.D. Goldberg, J. J. Gurka, Retrieving clear-sky atmospheric parameters from SEVIRI and ABI infrared radiances, Journal of Geophysical Research: Atmospheres, Aug. 2008, 113(D15). <https://doi.org/10.1029/2008JD010040>.
- [3] R.R. Irish, J.L. Barker, S.N. Goward, T. Arvidson, Characterization of the Landsat-7 ETM2 Automated Cloud-Cover Assessment (ACCA) Algorithm, American Society for Photogrammetry and Remote Sensing, 2006, pp. 1179-1188(10). <https://doi.org/10.14358/PERS.72.10.1179>.
- [4] L. Zhu, M. Wang, J. Shao, C. Liu, C. Zhao, Y. Zhao, Remote sensing of global volcanic eruptions using Fengyun series satellites, IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, Italy, Jul. (2015) 4797-4800. <https://doi.org/10.1109/IGARSS.2015.7326903>.
- [5] Z. Zhu and C. E. Woodcock, Object-based cloud and cloud shadow detection in Landsat imagery, Remote Sensing of Environment 118 (2012) 83-94. <https://doi.org/10.1016/j.rse.2011.10.028>.
- [6] S. Qiu, B. He, Z. Zhu, Z. Liao, X. Quan, Improving Fmask cloud and cloud shadow detection in mountainous area for Landsats 4-8 images, Remote Sensing of Environment 199 (2017) 107-119. <https://doi.org/10.1016/j.rse.2017.07.002>.
- [7] Y. Zhang, B. Guindon, J. Cihlar, An image transform to characterize and compensate for spatial variations in thin cloud contamination of Landsat images, Remote Sensing of Environment 82(2-3) (2002) 173-187. [https://doi.org/10.1016/S0034-4257\(02\)00034-2](https://doi.org/10.1016/S0034-4257(02)00034-2).
- [8] F. Xie, M. Shi, Z. Shi, J. Yin, D. Zhao, Multilevel Cloud Detection in Remote Sensing Images Based on Deep Learning, IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing 10(8) (2017) 3631-3640. <https://doi.org/10.1109/JSTARS.2017.2686488>.
- [9] J.H. Jeppesen, R.H. Jacobsen, F. Inceoglu, T.S. Toftgaard, A cloud detection algorithm for satellite imagery based on deep learning, Remote Sensing of Environment 229 (2019) 247-259. <https://doi.org/10.1016/j.rse.2019.03.039>.
- [10] W. Huang, Y. Wang, X. Chen, Cloud detection for high-resolution remote-sensing images of urban areas using colour and edge features based on dual-colour models, International Journal of Remote Sensing 39(20(2018) 6657-6675. <https://doi.org/10.1080/01431161.2018.1466069>.
- [11] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, O. Ronneberger, 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation, International conference on medical image computing and computer-assisted intervention, Springer, Cham, 2016 (424-432) [https://doi.org/10.1007/978-3-319-46723-8\\_49](https://doi.org/10.1007/978-3-319-46723-8_49).
- [12] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation, International Conference on Medical image computing and computer-assisted intervention, Springer, Cham 2015(234-241) [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28).
- [13] Z. Zhang, Q. Liu, Y. Wang, Road Extraction by Deep Residual U-Net, IEEE Geosci. Remote Sensing Letter 15(5)(2018) 749-753. <https://doi.org/10.1109/LGRS.2018.2802944>.
- [14] S. Mohajerani, T. A. Krammer, P. Saeedi, A Cloud Detection Algorithm for Remote Sensing Images Using Fully Convolutional Neural Networks, IEEE 20th International Workshop on Multimedia Signal Processing (MMSp), Vancouver, BC, Aug. 2018 (1-5) <https://doi.org/10.1109/MMSp.2018.8547095>.
- [15] B. Bischke, P. Helber, J. Folz, D. Borth, A. Dengel, Multi-Task Learning for Segmentation of Building Footprints with Deep Neural Networks, IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, Sep. (2019) 1480-1484. <https://doi.org/10.1109/ICIP.2019.8803050>.
- [16] S. Ji, S. Wei, M. Lu, Fully Convolutional Networks for Multisource Building Extraction From an Open Aerial and Satellite Imagery Data Set, IEEE Trans. Geosci. Remote Sensing 57(1) (2019) 574-586. <https://doi.org/10.1109/TGRS.2018.2858817>.
- [17] S. Mohajerani, P. Saeedi, Cloud-Net: An End-To-End Cloud Detection Algorithm for Landsat 8 Imagery, IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 2019 (1029-1032) <https://doi.org/10.1109/IGARSS.2019.8898776>.