



## Nghiên cứu quy trình tự động phát hiện biến động địa hình, đối tượng địa lý trên nền tảng công nghệ điện toán đám mây

Phạm Xuân Hoàn<sup>1\*</sup>, Phạm Hữu Hoàng<sup>1</sup>

Địa chỉ: Cục bản đồ/BTTM, số 2, ngõ 198 Trần Cung, Phường Cổ Nhuế 1, Quận Bắc Từ Liêm, Hà Nội

Email tác giả liên hệ: [hoanpx@gmail.com](mailto:hoanpx@gmail.com)

<https://doi.org/10.5281/zenodo.13845985>

### Tóm tắt:

Phương pháp tự động phát hiện biến động địa hình, địa vật trên bề mặt trái đất là vấn đề được đề cập và nghiên cứu từ lâu trong lĩnh vực Viễn thám nói riêng và khoa học trái đất nói chung. Các nghiên cứu gần đây đều tập trung vào việc nâng cao khả năng tự động phát hiện, giải đoán, phát hiện biến động các loại đối tượng trên ảnh Viễn thám ứng dụng công nghệ điện toán đám mây và trí tuệ nhân tạo. Hệ thống Google Earth Engine (GEE) là một nền tảng ứng dụng điện toán đám mây được xây dựng sẵn và miễn phí, cho phép tận dụng nguồn dữ liệu miễn phí phong phú, tận dụng hệ thống tính toán online công suất lớn, một số thuật toán được xây dựng sẵn để thực hiện một số bài toán Viễn thám phổ biến hiện nay. Trong giới hạn của bài báo này, nhóm tác giả sẽ giới thiệu quy trình tự động phát hiện biến động địa hình, đối tượng địa lý trên nền tảng công nghệ Google Earth Engine trên khu vực Hà Nội.

**Từ khóa:** Change Detection, GEE, Tự động phát hiện thay đổi, Randomforest

Ngày nhận bài: 08/07/2024

Ngày sửa lại: 16/07/2024

Ngày chấp nhận đăng: 18/07/2024

Ngày xuất bản: 30/09/2024

## Research on the process of automatically detecting terrain changes and geographic objects based on cloud computing technology

Pham Xuan Hoan<sup>1\*</sup>, Pham Huu Hoang<sup>1</sup>

<sup>1</sup> Defense Mapping Agency of VietNam, Ha Noi, Viet Nam

Corresponding Author Email: [hoanpx@gmail.com](mailto:hoanpx@gmail.com)

### Abstract:

The automatic method for detecting changes in terrain and surface features on the Earth's surface has been a topic of discussion and research for a long time in the field of Remote Sensing in particular and Earth Sciences in general. Recent studies have focused on enhancing the capability to automatically detect, interpret, and identify changes in various objects on remote sensing images using cloud computing and artificial intelligence technology. The Google Earth Engine (GEE) is a free cloud computing platform that allows users to leverage a rich source of free data, utilize a high-capacity online computing system, and use some pre-built algorithms to solve common remote sensing problems today. Within the scope of this paper, the authors will introduce an automated process for detecting changes in terrain and geographical objects using the Google Earth Engine technology platform in the Hanoi area.

**Keywords:** Change detection, GEE, Automatic change detection, Randomforest

Submission received: 08/07/2024

Revised: 16/07/2024

Accepted: 18/07/2024

Published: 30/09/2024

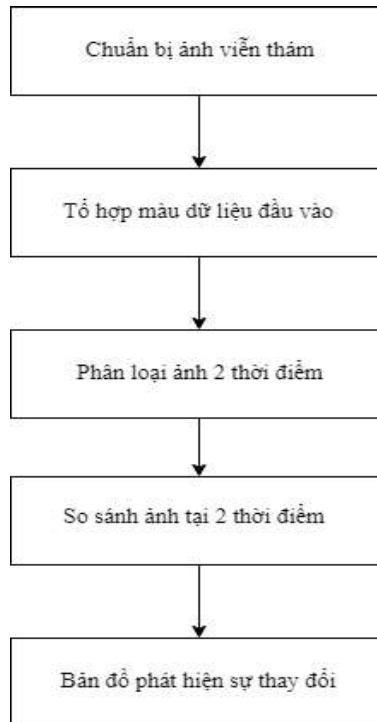
## 1. Mở đầu

Những năm gần đây, cùng với sự phát triển mạnh mẽ của kinh tế - xã hội thì sự biến động trên bề mặt lãnh thổ là rất lớn, nhiều hoạt động phá hủy tài nguyên thiên nhiên, cũng như đô thị hóa, các nhà máy xí nghiệp, khu công nghiệp mọc lên càng nhiều. Vì vậy việc phát hiện biến động của địa hình và các đối tượng địa lý góp phần đưa ra cảnh báo các khu có nhiều sự thay đổi để kịp thời khắc phục và đưa ra các giải pháp ứng phó [1].

Gần đây, công nghệ viễn thám được phát triển nhanh chóng với dữ liệu lớn và công nghệ điện toán đám mây. Google Earth Engine (GEE) là một nền tảng điện toán đám mây được phát triển để xử lý ảnh vệ tinh và các dữ liệu địa không gian khác. Hệ thống cung cấp quyền truy cập vào cơ sở dữ liệu khổng lồ ảnh vệ tinh và các thuật toán cần thiết để phân tích ảnh vệ tinh đáp ứng nhiều ứng dụng khác nhau. GEE cho phép giám sát những thay đổi trong các lĩnh vực nông nghiệp, tài nguyên nước và khí hậu... sử dụng dữ liệu địa không gian với các mức độ phân giải theo không gian và thời gian khác nhau. GEE cung cấp một danh mục dữ liệu cùng với thuật toán phân tích cho đa dạng nhóm người dùng [2, 3]. Việc sử dụng GEE đã tăng nhanh trong vài năm qua, trong nhiều cơ quan (nghiên cứu, chính phủ, và tư nhân) cũng như trong nhiều lĩnh vực ứng dụng khác nhau (quản lý nước, nông nghiệp, bảo tồn thiên nhiên, v.v.). Sự phổ biến của công cụ có liên quan đến những lợi ích to lớn của nó so với các công nghệ thông thường khác.

Ngoài ra, GEE là một nền tảng đám mây mạnh mẽ để xử lý và phân tích dữ liệu ảnh viễn thám, dưới đây là một số ứng dụng tiêu biểu của GEE [4, 5]:





Hình 2. Phương pháp giải đoán ảnh viễn thám, phát hiện thay đổi trên nền tảng GEE

### 2.3. Quy trình phân loại ảnh

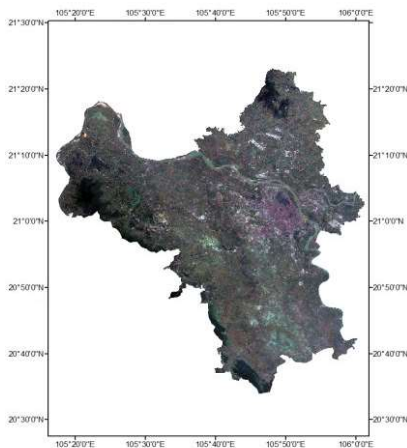
#### Bước 1. Chuẩn bị ảnh viễn thám

Trước khi bắt đầu quy trình phát hiện thay đổi ta cần chuẩn bị dữ liệu ảnh viễn thám để xử lý. Các bước này bao gồm đăng ký ảnh đa thời gian và hiệu chỉnh bức xạ và khí quyển, những bước này đặc biệt quan trọng. Tuy nhiên, phần lớn công việc này đã được tự động hóa và áp dụng cho các dữ liệu ảnh có sẵn trong GEE. Lựa chọn dữ liệu ảnh cho quá trình xử lý cũng rất quan trọng trong đó có việc tìm kiếm các ảnh có độ che phủ mây thấp.

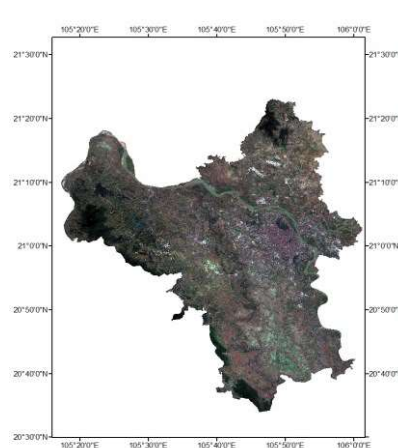
Lựa chọn ảnh Sentinel-2 ở hai thời điểm cần phân tích, so sánh sự biến động về địa hình, địa vật, trong nghiên cứu này chúng tôi chọn hai thời điểm trong vòng hai năm 2020 và 2021.

Tiếp theo, chọn bộ dữ liệu ảnh Sentinel-2 dựa trên tập lệnh ImageCollection của GEE tại 2 thời điểm tiến hành lọc nhiễu để có hình ảnh được chất lượng tốt nhất. Trong nghiên cứu này, tiến hành lọc tham số “CLOUDY\_PIXEL\_PERCENTAGE” nhỏ hơn 5% để thu được ảnh ít mây nhất (dưới 5%). Ngoài ra, còn áp dụng kỹ thuật Cloud masking để loại bỏ mây.

#### Bước 2. Tổ hợp màu tự nhiên dữ liệu ảnh đầu vào



Năm 2020



Năm 2021

Hình 3. Dữ liệu ảnh thu thập ở hai thời điểm



Hình 4. Lấy mẫu phân loại

Trước khi tiến hành phân tích phát hiện thay đổi, trước tiên chúng ta nên trực quan hóa hình ảnh đầu vào của mình để hiểu được khu vực nghiên cứu, kiểm tra trực quan nơi có thể xảy ra thay đổi và xác định bất kỳ vấn đề nào của dữ liệu đầu vào trước khi tiếp tục. Bằng cách tổ hợp màu kênh màu đỏ, xanh lá cây và xanh lam. Với ảnh Sentinel, tổ hợp màu tự nhiên bằng cách sử dụng Band 4 - kênh màu đỏ (Red), Band 3 - kênh màu xanh lá cây (Green) và Band 2 - kênh màu xanh lam (Blue).

Kết quả tổ hợp màu tự nhiên thu được ảnh tự nhiên tại 2 thời điểm

### Bước 3. Phân loại ảnh tại 2 thời điểm

Trong bài báo này chúng tôi tiến hành phân loại các đối tượng thực vật, thủy hệ, giao thông, dân cư. Sử dụng phương pháp phân loại có giám sát với các bước tiến hành như sau:

1. Chọn mẫu cho từng đối tượng giám sát

2. Sử dụng Machine Learning để GEE tiến hành phân loại các đối tượng trên tại 2 thời điểm, trong nghiên cứu này sử dụng thuật toán “smileRandomForest” hỗ trợ bởi GEE. Tiến hành đánh giá độ chính xác sau phân loại để kiểm chứng độ chính xác của quá trình phân loại có phù hợp để tiến hành so sánh ở bước tiếp theo.

### Bước 4. So sánh ảnh tại 2 thời điểm phát hiện đổi

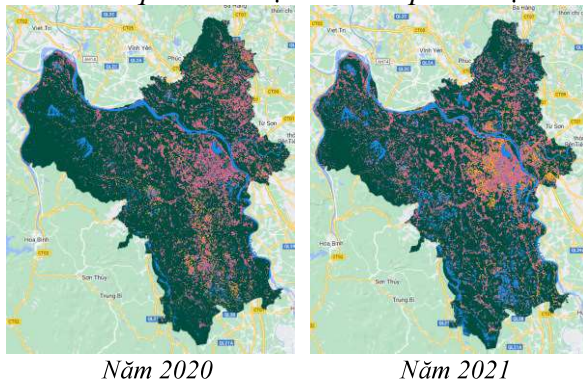
Sau khi tiến hành phân loại ảnh tại 2 thời điểm ta sử dụng thuật toán so sánh được hỗ trợ trong GEE để phát hiện sự thay đổi, lấy ảnh phân loại thời điểm sau loại trừ cho thời điểm trước. Gán màu cho các đối tượng nghiên cứu, để phát hiện vùng thay đổi, sau đó hiển thị lên bản đồ.

## 3. Kết quả nghiên cứu

- Dữ liệu sử dụng: Sentinel-2 đã được hiệu chỉnh khí quyển trong GEE.

- Phạm vi nghiên cứu: Khu vực nghiên cứu là toàn bộ diện tích Hà Nội.

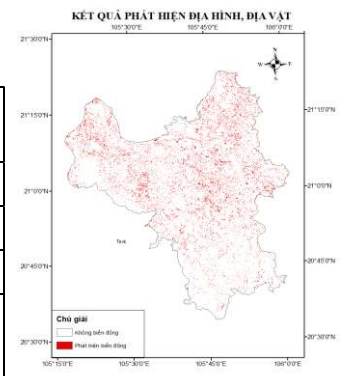
- Kết quả thu được sau khi phân loại khu vực Hà Nội được thể hiện tại Hình 5.



Hình 5. Ảnh sau phân loại ở hai thời điểm

Bảng 1. Màu sắc của các lớp phủ

| STT | Đối tượng  | Màu sắc mẫu |
|-----|------------|-------------|
| 1   | Thực vật   |             |
| 2   | Thủy hệ    |             |
| 3   | Dân cư     |             |
| 4   | Giao thông |             |



Hình 6. Kết quả phát hiện biến động địa hình, địa vật

### 3.1. Đánh giá độ chính xác kết quả phân loại

Phân loại ảnh viễn thám là quá trình phân định các pixel trong hình ảnh thành các lớp hoặc các nhóm đơn vị lớp phủ bề mặt đất [6-7]. Để thực hiện chúng ta cần có bộ khóa giải đoán và bộ khóa đánh giá độ chính xác. Việc chọn mẫu để đánh giá độ chính xác được thực hiện qua xác định vị trí các điểm mẫu ở thực địa để so sánh và đánh giá với kết quả giải đoán ảnh. Thực hiện với dung lượng mẫu tương đối phù hợp với diện tích giải đoán để có được sự tổng quát khách quan và bao quát các đối tượng cần đánh giá sau đó chồng xếp lên lớp bản đồ nền bản đồ giải đoán để kiểm tra độ chính xác của kết quả.

Thực hiện so sánh kết quả giải đoán và kết quả thực địa để thành lập ma trận sai số (Confusion Matrix) [8-9]. Ma trận sai số là một bản ma trận thể hiện sự sai khác và trùng khớp kết quả kiểm tra thực địa và kết quả giải đoán. Từ đó tính được chỉ số Kappa (K) để đánh giá độ chính xác của kết quả giải đoán. Hệ số Kappa được sử dụng là thước đo đánh giá độ chính xác phân loại. Đây là hệ số tiện ích của tất cả các nguyên tố từ ma trận sai số. Nó là sự khác nhau cơ bản giữa những gì có thực về sai số độ lệch của ma trận và tổng số thay đổi được chỉ ra bởi hàng và cột.

Công thức xác định chỉ số Kappa như sau [10]:



$$\frac{N \sum_{i=1}^r X_{ii} - \sum_{i=1}^r (X_{i+} - X_{+i})}{N^2 - \sum_{i=1}^r (X_{i+} - X_{+i})} \quad (1)$$

Trong đó:  $r$  = Số lượng cột trong ma trận ảnh,  $X_{ii}$  = Số lượng Pixel quan sát được tại hàng  $i$  và cột  $i$  (trên đường chéo chính)  $X_{i+}$  = Tổng pixel quan sát tại hàng  $i$ ,  $X_{+i}$  = Tổng pixel quan sát tại cột  $i$ , và  $N$  = Tổng số pixel quan sát được trong ma trận ảnh.

Hệ số Kappa thường nằm giữa 0 và 1, giá trị nằm trong khoảng này thì độ chính xác của sự phân loại được chấp nhận. Theo Cục Địa chất Mỹ, Kappa có 3 nhóm giá trị:  $K > 0,8$ : độ chính xác cao,  $0,4 < K < 0,8$ : độ chính xác vừa phải, và  $K < 0,4$ : độ chính xác thấp

Bảng 2. Ma trận Confusion của ảnh năm 2020

| Lớp                     | Đô thị | Đất trồng | Thủy hệ | Thực vật | Tổng  | ĐCX người dùng                   |
|-------------------------|--------|-----------|---------|----------|-------|----------------------------------|
| Đô thị                  | 433    | 5         | 0       | 3        | 441   | 98,18                            |
| Giao thông              | 2      | 222       | 0       | 7        | 231   | 96,1                             |
| Thủy hệ                 | 0      | 0         | 15783   | 11       | 15794 | 99,93                            |
| Thực vật                | 0      | 0         | 1       | 10507    | 10508 | 99,99                            |
| <b>Tổng</b>             | 435    | 227       | 15784   | 10528    |       |                                  |
| <b>ĐCX nhà sản xuất</b> | 99,54  | 97,79     | 99,99   | 99,8     |       | - ĐCX: 99,89%<br>- Kappa: 0.9978 |

Bảng 3. Ma trận Confusion của ảnh năm 2021

| Lớp                     | Đô thị | Đất trồng | Thủy hệ | Thực vật | Tổng  | ĐCX người dùng                   |
|-------------------------|--------|-----------|---------|----------|-------|----------------------------------|
| Đô thị                  | 433    | 5         | 0       | 3        | 441   | 98,18                            |
| Giao thông              | 2      | 222       | 0       | 7        | 231   | 96,1                             |
| Thủy hệ                 | 0      | 0         | 15783   | 11       | 15794 | 99,93                            |
| Thực vật                | 0      | 0         | 1       | 10507    | 10508 | 99,99                            |
| <b>Tổng</b>             | 435    | 227       | 15784   | 10528    |       |                                  |
| <b>ĐCX nhà sản xuất</b> | 99,54  | 97,79     | 99,99   | 99,8     |       | - ĐCX: 99,89%<br>- Kappa: 0.9978 |

Dựa trên các kết quả từ Ma trận nhầm lẫn tại các thời điểm nêu trên, nghiên cứu đã chỉ ra rằng độ chính xác đạt mức độ cao, với nghiên cứu khoa học có thể chấp nhận được.

### 3.2. Kết quả biến động trên bề mặt

- Những vùng màu trắng: Khu vực không bị thay đổi
- Vùng màu đỏ: Khu vực có sự biến động của các đối tượng
- Thời gian xử lý: Dưới 3 phút.

Bảng 7. Kết quả các khu vực biến động

| TT | Đối tượng             | Diện tích (m <sup>2</sup> ) | TT | Đối tượng             | Diện tích (m <sup>2</sup> ) |
|----|-----------------------|-----------------------------|----|-----------------------|-----------------------------|
| 1  | Đô thị → Giao thông   | 731486                      | 7  | Thủy hệ → Đô thị      | 32348                       |
| 2  | Đô thị → Thủy hệ      | 25338                       | 8  | Thủy hệ → Giao thông  | 213411                      |
| 3  | Đô thị → Thực vật     | 2048116                     | 9  | Thủy hệ → Thực vật    | 1274626                     |
| 4  | Giao thông → Đô thị   | 264406                      | 10 | Thực vật → Đô thị     | 830485                      |
| 5  | Giao thông → Thủy hệ  | 17485                       | 11 | Thực vật → Giao thông | 2094208                     |
| 6  | Giao thông → Thực vật | 628903                      | 12 | Thực vật → Thủy hệ    | 227256                      |

Qua bản đồ biến động kết quả phát hiện sự thay đổi ở một số vị trí như sau:

Ở khu vực Trung tâm thấy có sự thay đổi nhiều nhất: Sự thay đổi ở sự biến động của khu đô thị (nhà, công trình được xây dựng thêm khá nhiều), giao thông.

Khu vực ít biến đổi nhất, Tây Bắc đến Tây Nam chủ yếu khu vực này là rừng, thực vật nên ít biến động.

Tuy nhiên, trong quá trình phát hiện sự thay đổi còn phụ thuộc vào độ chính xác của quá trình chọn mẫu phân loại của ảnh trước và sau. Loại ảnh sử dụng cũng cho chúng ta một số kết quả khác nhau tương đối, do vậy cần phải lựa chọn các loại ảnh tương đối ổn định, ít bị che phủ bởi mây và ảnh hưởng bởi thời tiết. Ngoài ra, việc lựa chọn thuật toán xử lý cũng ảnh hưởng nhiều đến kết quả tính toán do đó cần phải lựa chọn các thuật toán phù hợp với từng vùng địa lý nhất định để đạt được kết quả chính xác.

## 4. Kết luận

Việc nghiên cứu quy trình tự động phát hiện biến động địa hình, đối tượng địa lý trên nền tảng công nghệ điện toán đám mây và AI là công nghệ mới để giải quyết nhanh bài toán phát hiện thay đổi trên diện rộng; giải đoán đối tượng địa lý. Nghiên cứu sẽ làm phong phú thêm nguồn dữ liệu đầu vào để xử lý và nâng cao khả năng tham mưu cho lãnh đạo, chỉ huy về phạm vi biến động nhiều để có thể ra quyết định trong việc hoạch định các kế hoạch về quy hoạch, quản lý trong đó có việc thành lập mới, cập nhật hoặc chỉnh lý bản đồ ở các khu vực khác nhau trong thời gian nhanh. Đặc biệt, chúng ta tận dụng được nguồn tư liệu viễn thám miễn phí độ phân giải cao, phạm vi rộng, độ



phủ độ chính xác đồng nhất để tiến hành phân tích, bỏ qua quá trình chờ mua ảnh thương mại vừa không hiệu quả về kinh tế và đôi khi không có tính thời sự.

Đối với Cục Bản đồ theo chu kỳ hàng năm hoặc kế hoạch đột xuất cần phải có sự đánh giá về biến đổi địa hình địa vật trên một phạm vi rộng (trên một quân khu, một miền, cả nước hoặc trên một khu vực bất kỳ) để Tham mưu cho Thủ trưởng Cục phạm vi biến động nhiều để có thể ra quyết định về việc thành lập mới, cập nhật hoặc chỉnh lý bản đồ ở các khu vực đó. Việc phát hiện thay đổi như trên là rất cần thiết để hoạch định kế hoạch cập nhật, chỉnh lý bản đồ các loại tỷ lệ ở các khu vực có biến động lớn. Đối với các quân, binh chủng khác việc phát hiện thay đổi, biến động địa hình, địa vật trên phạm vi rộng cũng là một thông tin vô cùng hữu ích trong các hoạt động sẵn sàng chiến đấu và chiến đấu của đơn vị.

#### Tài liệu tham khảo

- [1] Tô Văn Tường, “Quy hoạch tài nguyên nước lưu vực sông Đồng Nai”, *Dự án cấp Bộ - Viện Quy hoạch Thủy lợi miền Nam*, 2009.
- [2] Viện Khoa học Thủy lợi miền Nam, “Rà soát quy hoạch thủy lợi vùng Đông Nam Bộ phục vụ tái cấu trúc ngành nông nghiệp”, *Dự án quy hoạch cấp*, 2015.
- [3] McFeeters S K, “The use of Normalized Difference Water Index (NDWI) in the delineation of open water features”. *International Journal of Remote sensing*, 1996.
- [4] N. Patel E A, “Multitemporal settlement and population mapping from Landsat using google earth engine”. *Earth Obs Geoinf*, 2015.
- [5] Nguyễn Hoàng Quân, “Một số phương pháp xây dựng bản đồ ngập lũ tỉnh Long An trong điều kiện biến đổi khí hậu, nước biển dâng”. *Science & Technology*, 2013.
- [6] Carrasco L, O’Neil A W, Morton R D, Rowland C S, “Evaluating Combinations of Temporally Aggregated Sentinel-1, Sentinel-2 and Landsat 8 for Land Cover Mapping with Google Earth Engine”, 2019.
- [7] Google Earth Engine API. Introduction, ngày truy cập 22/12/2023. Địa chỉ <https://developers.google.com/earth-engine>
- [8] Nguyễn Văn Hoàng, “Nghiên cứu ứng dụng nền tảng Google Earth Engine thành lập bản đồ giám sát hạn hán lưu vực sông Đồng Nai vùng Đông Nam Bộ”, *Tạp chí khoa học và công nghệ đại học thủy lợi*, số 58, 2020.
- [9] Trịnh Minh Ngọc, Nguyễn Quang Hưng, “Ứng dụng nền tảng Google Earth Engine xây dựng bản đồ phân vùng hạn hán cho tỉnh Bình Thuận”, *Tạp chí khoa học biến đổi khí hậu* số 24 – Tháng 12/2022.
- [10] Đinh Bảo Ngọc, Nông Thị Oanh, Phạm Quang, “Nghiên cứu nền tảng Google Earth Engine phục vụ xây dựng các hệ thống xác định vị trí sạt lở đất từ dữ liệu viễn thám đa thời gian”, *Tạp chí khoa học Đo đạc và bản đồ*, số 57-9/23, 2023.

