

Vector, Raster và bài toán dự đoán Land-use, Land-cover

{ Nguyễn Văn Chọn }¹

¹ Trường Đại học Công nghệ Thông tin, Thủ Đức, TP.Hồ Chí Minh

Đặt vấn đề. Hiện nay hệ thống thông tin địa lý đang được áp dụng vào nhiều lĩnh vực đặc biệt là trong GIS (geographic information systems). Có 2 loại dữ liệu thường được sử dụng xuyên suốt đối với các nhà nghiên cứu về ứng dụng GIS 3D đó là dạng dữ liệu vector và dạng dữ liệu raster. Trong bài nghiên cứu này chúng tôi tiến hành phân tích chi tiết về 2 loại dữ liệu trên, sau đó tiến hành thực nghiệm trong bài toán land use – land cover. Bước đầu cho thấy dữ liệu vector cho ra kích thước file nhẹ và tốc độ tính toán nhanh hơn, biểu diễn chính xác về ranh giới hành chính hơn. Tuy nhiên lại không thể biểu diễn các chi tiết phức tạp mà phải cần đến dữ liệu raster. Sau cùng, chúng tôi cho vào mô hình học máy để so sánh độ chính xác giữa hai bộ dữ liệu. Kết quả dữ liệu raster cho hiệu quả dự đoán cao hơn với 90% ở mô hình KNN. Chúng tôi cũng đang phát triển kết hợp 2 dạng dữ liệu để tạo ra mô hình hybrid dễ dàng biểu diễn các thuộc tính tự nhiên hơn trong tương lai.

Từ khoá: Vector, Raster, Machine Learning, Land-use, Land-cover,...

1 Giới thiệu

1.1 Land-use

Sử dụng đất (Land-use) là khái niệm được xây dựng dựa theo chức năng, mục đích sử dụng đất. Do vậy, một lớp sử dụng đất có thể được định nghĩa là một tập hợp các hành động được thực hiện nhằm cung cấp một hay nhiều hơn loại hàng hóa hoặc dịch vụ. Một lớp sử dụng đất cho trước có thể có ở nhiều mảnh đất khác nhau hoặc trên một mảnh đất có thể có nhiều hành động sử dụng đất khác nhau. Loại đất cho thấy cách con người sử dụng cảnh quan - cho mục đích phát triển, bảo tồn hay sử dụng hỗn hợp.

Land use có ảnh hưởng lớn đến môi trường, phát triển kinh tế, xã hội, và đời sống cộng đồng. Một số mục đích phổ biến của việc sử dụng đất: Đô thị và Kinh doanh, Nông nghiệp, bảo tồn tự nhiên và đất rừng, công nghiệp, hạ tầng giao thông

1.2 Land-cover

Lớp phủ mặt đất (land-cover) là lớp phủ vật chất quan sát được khi nhìn từ mặt đất hoặc thông qua vệ tinh viễn thám, bao gồm thực vật (mọc tự nhiên hoặc được trồng cây) và các cơ sở xây dựng của con người (nhà cửa, đường sá...) bao phủ bề mặt mặt đất. Nước, băng, đá lở hay các bãi cát cũng được coi là lớp phủ mặt đất.

Độ che phủ đất có thể được xác định bằng cách phân tích hình ảnh vệ tinh và trên không. Việc sử dụng đất không thể được xác định từ hình ảnh vệ tinh. Để thấy sự thay đổi theo thời gian, cần có bản đồ che phủ đất trong nhiều năm khác nhau.

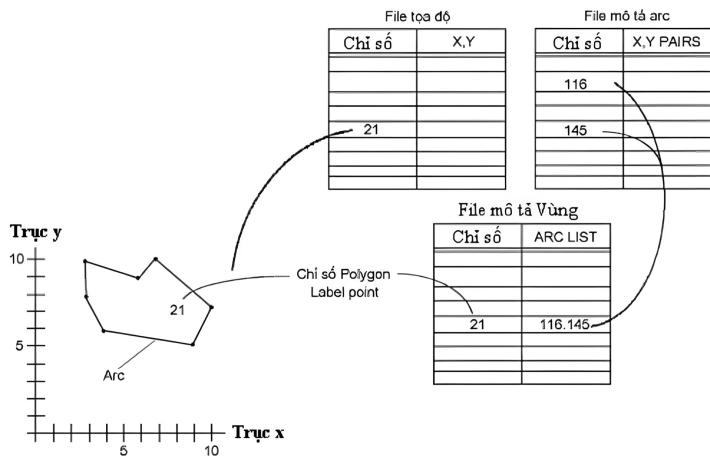
2 Thực nghiệm

2.1 Các đặc trưng của vector

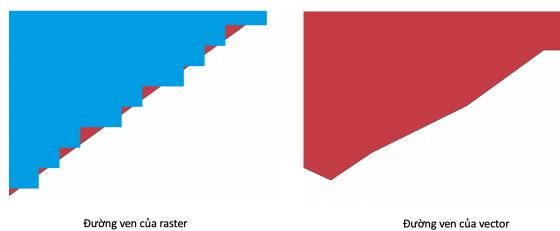
Phương pháp vector có khả năng biểu diễn topology (tính chất không gian) một cách rất mạnh mẽ. Dữ liệu vector biểu diễn không gian bằng cách sử dụng các đối tượng hình học như điểm, đường, và vùng. Các yếu tố topology như mối quan hệ đỉnh (nodes), cạnh (edges), và mặt (faces) hoặc vùng (polygons) giữa các đối tượng được bảo toàn và có thể được mô tả chính xác.

Biểu diễn đường ven của vector có thể nói là ưu việt hơn hẳn so với raster khi vector, biểu diễn thường là các cạnh thẳng liền kề bao sát lấy phần bì mặt bên ngoài của khu vực. Các đường mảnh của vector có thể giảm thiểu sự chênh lệch giữa khu vực thực và khu vực được cắt đi cho các phép tính của vector.

Trong nhiều trường hợp dữ liệu raster có thể bị tính toán sai do thành phần cấu tạo của raster được tạo từ các điểm pixel để biểu diễn nhiều thuộc tính trong một phạm vi nhất định. Các ô hình vuông hoặc hình chữ nhật sẽ là cho việc tính toán trở nên chậm hơn. Và độ chính xác cũng trở nên kém hơn khi giữa các đường ven bao quanh có một số khoảng trống không cần thiết hoặc các khoảng lồi do khói hình tạo ra.

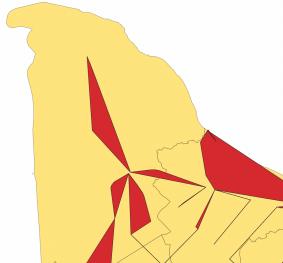


Hình 1 Biểu diễn vùng theo dữ liệu vector



Hình 2 So sánh đường ven của raster, vector

Ngoài ra Vector còn cung cấp các phương thức để có thể dễ dàng thực hiện tính toán trên bản đồ tương ứng. Đặc biệt là đối với các công trình, tòa nhà hoặc đường xá. Ví dụ bên dưới mô tả về đường bộ ở các tỉnh thành, chúng tôi sử dụng các đường bao vector để chỉ ra những điểm ảnh hưởng của đường xá đối với người dân sống ở gần đó. Có nghĩa là người dân ở trong khu vực có thể dễ dàng tiếp cận đến tuyến đường giao thông này. Việc này có thể giúp chúng ta có cái nhìn bao quát việc tiếp cận đường xá của người dân trong khu vực từ đó đưa ra phương hướng, ưu tiên xây dựng các tuyến đường quan trọng cho người dân ở một số khu vực nhất định.



Hình 3 Bao lồi vector đường xá

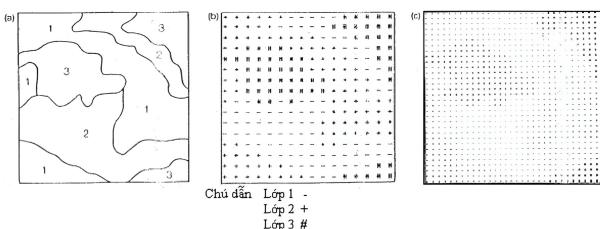
2.2 Các đặc trưng của raster

Trong raster, dữ liệu được biểu diễn bằng lưới các ô vuông (pixel), và mỗi pixel chứa một giá trị tại một vị trí cụ thể. Raster không giữ thông tin về mối quan hệ không gian giữa các pixel, do đó, không thể biểu diễn chính xác các yếu tố topology như đỉnh, cạnh, hay các đối tượng không gian phức tạp.

Phương pháp raster thích hợp cho việc biểu diễn dữ liệu liên tục như độ cao, màu sắc, nhiệt độ, và các thông tin không gian khác, nhưng không phải là lựa chọn tốt nhất nếu yêu cầu biểu diễn chính xác các đối tượng có tính chất topology phức tạp. Đối với các nhiệm vụ yêu cầu topology, phương pháp vector thường được ưu chuộng hơn.

Một số đặc điểm nổi bật phải kể đến như:

- Các điểm được xếp liên tiếp từ trái qua phải và từ trên xuống dưới.
- Mỗi một điểm ảnh (pixcel) chứa một giá trị.
- Một tập các ma trận điểm và các giá trị tương ứng tạo thành một lớp (layer).
- Trong cơ sở dữ liệu có thể có nhiều lớp.



Hình 4. Các lớp raster có thuộc tính khác nhau có thể chồng chập

Trong biểu diễn raster, mỗi pixel có thể được gắn với giá trị độ cao tương ứng. Điều này rất hữu ích trong việc biểu diễn đồng đều và chính

xác chiều cao của các vùng địa lý. Đối với dữ liệu độ cao, raster thường được sử dụng trong các mô hình địa lý, như mô hình đồng đều độ cao (DEM - Digital Elevation Model) [1].

Tuy nhiên cũng có một số lỗi khi sử dụng raster cho việc đo độ cao của các ngọn núi trong tự nhiên. Khi ta sử dụng chúng để đo đạc hoặc biểu diễn chúng dưới dạng độ dốc thì raster lại chỉ ra cho mặt phẳng giống dạng 2D [2].



Hình 5. Lược đồ cao nguyên ở Tây Nguyên

Ưu điểm của nó sẽ dễ dàng cho việc biểu diễn các cao nguyên rộng lớn, độ dốc thấp không nhìn rõ bằng mắt thường. Nhưng ngược lại ở những hẻm núi sâu, nhọn thì việc biểu diễn độ dốc và độ cao cần phải có sự có mặt của vector để nhìn trực quan hơn.

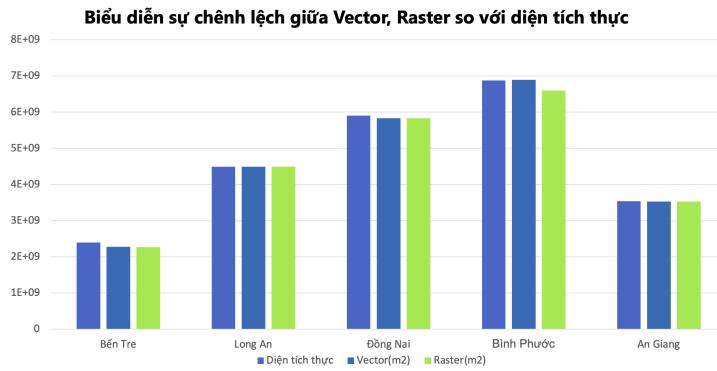
Hơn nữa, đối với một số công trình phức tạp có nhiều thuộc tính. Hoặc có những nhiều sự vật cần phải biểu thị thì lúc này chúng ta không thể sử dụng vector để tính toán và biểu diễn được. Mặc dù Vector biểu diễn rất chính xác các đối tượng cụ thể và các công trình, nhưng khi biểu diễn chúng với một số lượng lớn thì việc tăng kích thước tệp là điều không thể tránh khỏi. Lúc này raster là lựa chọn hiệu quả hơn.



Hình 6. Biểu diễn công trình phức tạp bằng raster, vector

2.3 Các thang đo

Chúng tôi tiến hành thực nghiệm đo diện tích trên 5 tỉnh thành có đa dạng các địa hình và nhiều sông hồ. Kết quả cho thấy diện tích của 2 bộ dữ liệu về vector và raster có sự chênh lệch ít.



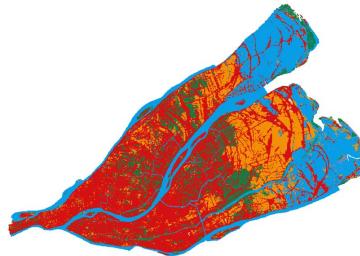
Hình 7 Biểu đồ so sánh diện tích của vector, raster, dữ liệu thực

Chúng ta có thể nhìn rõ sự chênh lệch hơn ở biểu đồ phía bên dưới:



Hình 8 Biểu đồ sự chênh lệch về độ chính xác khi tính toán diện tích của vector và raster

Khi tính toán lại dữ liệu thì ta có thể thấy độ chênh lệch của tỉnh Bến Tre và tỉnh Đồng Nai là lớn nhất. Chúng tôi đã tìm hiểu nguyên nhân thì thấy ở 2 tỉnh này thường có nhiều sông ngòi và ao hồ, việc này không chỉ làm ảnh hưởng đến độ chính xác khi tính toán diện tích mà còn làm thay đổi thể tích bề mặt một cách đáng kể.



Hình 9 Dữ liệu raster tỉnh Bến Tre

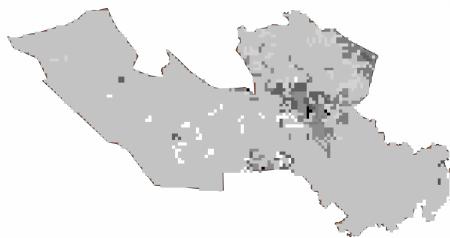
Ngoài ra chúng tôi còn tiến hành đo đạc về thời gian thực thi khi tính diện tích của từng tỉnh. Kết quả cho thấy việc tính toán dữ liệu bằng raster sẽ mất nhiều thời gian hơn so với dữ liệu bằng vector khoảng 200 lần.

Hơn nữa kích thước file dữ liệu của vector cũng rất nhỏ. Một trong số những nguyên nhân trên có thể kể đến là raster có nhiều lớp dữ liệu hơn, nên việc tính toán và kích thước dữ liệu nhiều hơn.

Tỉnh	Vector (S)	Raster (S)	Diện tích thực	Vector (m2)	Raster (m2)	Vector (KB)	Raster(MB)
Bến Tre	0.003661	0.74	2394000000	2272782661	2272409600	148	47.75
Long An	0.003854	1.01	4493800000	4492649094	4490995500	212	91.65
Đồng Nai	0.003860	0.97	5903400000	5832149291	5831762100	346	95.52
Bình Phước	0.004280	1.12	6873560000	6893162865	6895132500	300	116.94
An Giang	0.003978	0.78	3536760000	3531599282	3528843200	137	71.68

Hình 10 Bảng đo đặc tổng quan

Thí nghiệm tiếp theo chúng tôi tiến hành gray scale dữ liệu raster và tính toán lại xem kết quả có thay đổi hay không. Chúng tôi đã sử dụng tỉnh Long An để tiếp tục đo thời gian tính diện tích. Kết quả có sự thay đổi rõ rệt, thời gian từ 1.01s chuyển xuống còn 0.4s, Việc giảm thuộc tính của bộ dữ liệu raster sẽ cho ra thời gian tính toán sớm hơn, nhưng so lại với dữ liệu vector là 0.0038 thì vẫn còn cách biệt rất rõ ràng.



Hình 11 Raster gray scale

2.4 Tổng quan sơ bộ

Ưu nhược điểm của dữ liệu vector:

Ưu điểm	Nhược điểm
Biểu diễn chính xác ranh giới hành chính	Khó mô tả chính xác các đối tượng tự nhiên
Kích thước dữ liệu nhỏ, xử lý nhanh	Khó tích hợp các thuộc tính không gian phức tạp
Thuận tiện cho biên tập và cập nhật dữ liệu	Phức tạp hóa các phân tích và mô hình hóa

Ưu nhược điểm của dữ liệu raster:

Ưu điểm	Nhược điểm
Dễ dàng tích hợp các thuộc tính không gian	Kích thước dữ liệu lớn, tốn bộ nhớ tính toán
Thuận tiện cho việc phân tích không gian và thống kê	Độ phân giải ảnh ảnh hưởng đến độ chính xác
Các thuật toán xử lý ảnh có thể áp dụng trực tiếp	Khó biểu diễn chính xác ranh giới hành chính

3 Áp dụng máy học vào dự đoán land-use, land-cover

3.1 Mục đích

Mục đích chính là nhận dạng và phân loại các loại đất sử dụng khác nhau (land-cover) như rừng, đồng cỏ, mặt nước, đất canh tác, khu dân cư,... từ các ảnh vệ tinh, cho phép theo dõi và giám sát các thay đổi sử dụng đất theo thời gian.

Kết quả dự đoán land-cover có thể được sử dụng trong nhiều ứng dụng khác nhau như quy hoạch đô thị, nông nghiệp, lâm nghiệp, quản lý tài nguyên, môi trường...

3.2 Bộ dữ liệu raster

a. Bộ dữ liệu phủ đất (land-cover):

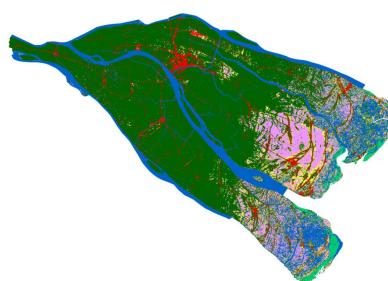
Gồm 11 lớp được xác định dựa trên Hệ thống Phân loại Phủ đất (Land Cover Classification System - LCCS) do Tổ chức Nông nghiệp và Lương thực Liên hiệp quốc (FAO) phát triển.

Hệ thống LCCS được thiết kế theo cấu trúc phân cấp, cho phép điều chỉnh chi tiết của bản đồ phù hợp với lượng thông tin có sẵn.

11 lớp gồm: Tree cover, Shrubland, Grassland, Cropland, Built-up, Bare ground, Snow and ice, Permanent waterbodies, Herbaceous wetland, Mangroves, Moss and lichen

Map code	Land Cover Class	LCCS code	Definition	Color code (RGB)
10	Tree cover	A12A3 // A11A1 A24A3C1(C2)-R1(R2)	This class includes any geographic area dominated by trees with a cover of 10% or more. Other land cover classes (shrubs and/or herbs in the understorey, built-up, permanent water bodies, ...) can be present below the canopy, even with a density higher than trees. Areas planted with trees for afforestation purposes and plantations (e.g. oil palm, olive trees) are included in this class. This class also includes tree covered areas seasonally or permanently flooded with fresh water except for mangroves.	0,100,0
20	Shrubland	A12A4 // A11A2	This class includes any geographic area dominated by natural shrubs having a cover of 10% or more. Shrubs are defined as woody perennial plants with persistent and woody stems and without any defined main stem being less than 5 m tall. Trees can be present in scattered form if their cover is less than 10%. Herbaceous plants can also be present at any density. The shrub foliage can be either evergreen or deciduous.	255, 187, 34
30	Grassland	A12A2	This class includes any geographic area dominated by natural herbaceous plants (Plants without persistent stem or shoots above ground and lacking definite firm structure): (grasslands, prairies, steppes, savannahs, pastures) with a cover of 10% or more, irrespective of different human and/or animal activities, such as: grazing, selective fire management etc. Woody plants (trees and/or shrubs) can be present assuming their cover is less than 10%. It may also contain uncultivated cropland areas (without harvest/ bare soil period) in the reference year	255, 255, 76
40	Cropland	A11A3(A4)(A5) // A23	Land covered with annual cropland that is sowed/planted and harvestable at least once within the 12 months after the sowing/planting date. The annual cropland produces an herbaceous cover and is sometimes combined with some tree or woody vegetation. Note that perennial woody crops will be classified as the appropriate tree cover or shrub land cover type. Greenhouses are considered as built-up.	240, 150, 255
50	Built-up	B15A1	Land covered by buildings, roads and other man-made structures such as railroads. Buildings include both residential and industrial building. Urban green (parks, sport facilities) is not included in this class. Waste dump deposits and extraction sites are considered as bare.	250, 0, 0
60	Bare / sparse vegetation	B16A1(A2) // B15A2	Lands with exposed soil, sand, or rocks and never has more than 10 % vegetated cover during any time of the year	180, 180, 180
70	Snow and ice	B28A2(A3)	This class includes any geographic area covered by snow or glaciers persistently	240, 240, 240
80	Permanent water bodies	B28A1(B1) // B27A1(B1)	This class includes any geographic area covered for most of the year (more than 9 months) by water bodies: lakes, reservoirs, and rivers. Can be either fresh or salt-water bodies. In some cases the water can be frozen for part of the year (less than 9 months).	0, 100, 200
90	Herbaceous wetland	A24A2	Land dominated by natural herbaceous vegetation (cover of 10% or more) that is permanently or regularly flooded by fresh, brackish or salt water. It excludes unvegetated sediment (see 60), swamp forests (classified as tree cover) and mangroves see 95).	0, 150, 160
95	Mangroves	A24A3C5-R3	Taxonomically diverse, salt-tolerant tree and other plant species which thrive in intertidal zones of sheltered tropical shores, "overwash" islands, and estuaries.	0, 207, 117
100	Moss and lichen	A12A7	Land covered with lichens and/or mosses. Lichens are composite organisms formed from the symbiotic association of fungi and algae. Mosses contain photo-autotrophic land plants without true leaves, stems, roots but with leaf-and stemlike organs.	250, 230, 160

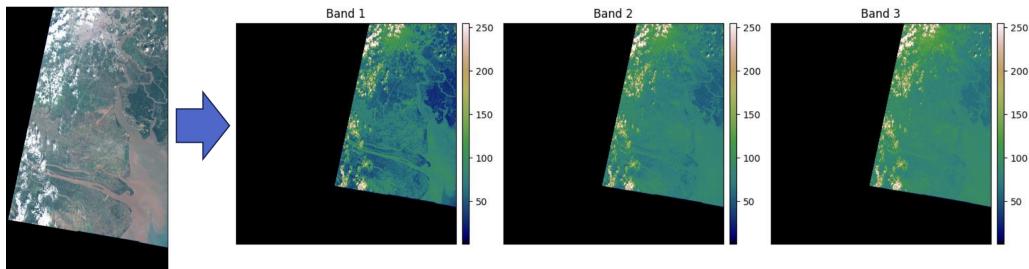
Hình ảnh minh họa ví dụ:



b. Bộ dữ liệu ảnh vệ tinh:

Gồm ảnh RGB bì mặt Trái Đất. Là ảnh đa phổ Sentinel-2 chụp bì mặt Trái đất và được chụp từ vệ tinh Sentinel-2 của Cơ quan Vũ trụ Châu Âu (ESA).

Dùng làm dữ liệu đầu vào để mô hình dự đoán các lớp phủ đất.

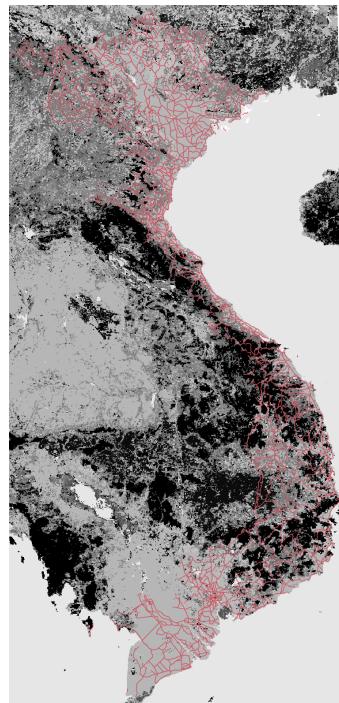


3.3 Bộ dữ liệu vector

Gồm 2 bộ dữ liệu, về tỉ lệ che phủ đất (format *vrt) và bản đồ đường giao thông (format *shp) phục vụ cho việc dự đoán tỉ lệ sử dụng đất.

Ảnh mô tả bộ dữ liệu gồm 2 lớp chồng nhau:

- Bản đồ che phủ đất (land-cover) làm lớp nền
- Bản đồ đường giao thông (màu đỏ) ở phía trên



3.4 Phương pháp xử lý

a. Dữ liệu raster

Kết quả dữ liệu đầu vào lớp land-cover sẽ được phân tích lấy các đoạn mã (code) dùng làm lớp mục tiêu (tổng 11 lớp).

Dữ liệu không gian đầu vào được phân tích biểu diễn dưới 3 hệ màu

(Blue, green, red) tương ứng với lớp land-cover phía trên.

Code	Blue	Green	Red
10	1923	1895	1702
10	1823	1741	1496
30	2443	2469	2636
50	3962	4023	4406
30	2195	2261	2315

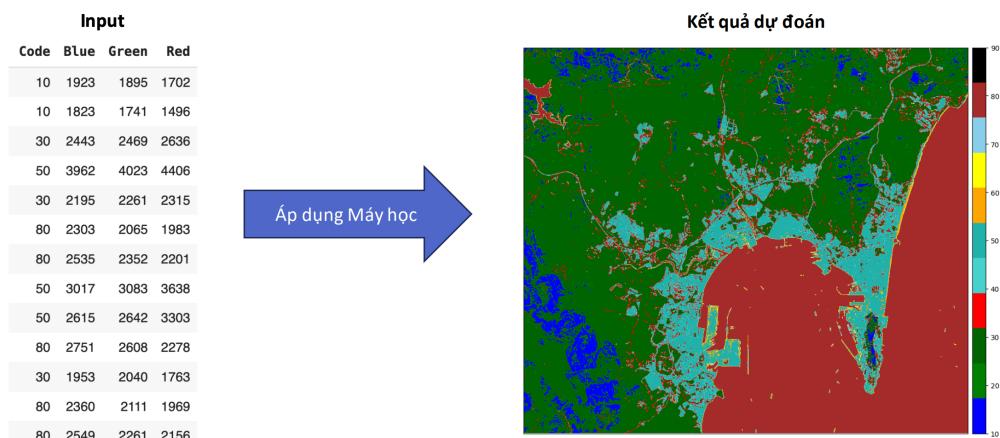
b. Dữ liệu vector

Tương ứng với 2 bộ dữ liệu ở trên (về land-cover và tuyến đường), bằng việc phân tích các tuyến đường, chúng ta sẽ tìm được ác cột start_x, start_y, end_x và end_y tương ứng. Các dữ liệu này sẽ được sử dụng cho training data.

	MED_DESCRI	RTT_DESCRI	F_CODE_DES	ISO	ISOCOUNTRY	index_right	land_use	start_x	start_y	end_x	end_y
0	None	None	Trail	VNM	VIETNAM	2969	16.0	105.363167	23.277250	105.298031	23.367803
3	None	None	Trail	VNM	VIETNAM	2969	16.0	105.073975	23.122000	104.954896	23.179873
12	Without Median	Secondary Route	Road	VNM	VIETNAM	2969	16.0	104.888010	23.171169	104.994942	23.065277
0	None	None	Trail	VNM	VIETNAM	1761	9.0	105.363167	23.277250	105.298031	23.367803
0	None	None	Trail	VNM	VIETNAM	3190	13.0	105.363167	23.277250	105.298031	23.367803
...
2090	Without Median	Secondary Route	Road	VNM	VIETNAM	88505	20.0	106.077141	9.733444	105.977448	9.610611
2115	Without Median	Secondary Route	Road	VNM	VIETNAM	88596	12.0	105.266357	9.508028	105.394165	9.378250
2135	Without Median	Secondary Route	Road	VNM	VIETNAM	88621	12.0	105.095055	9.340361	105.144722	9.177944
2140	Without Median	Secondary Route	Road	VNM	VIETNAM	88655	20.0	105.150803	9.173667	104.995228	8.757629
2140	Without Median	Secondary Route	Road	VNM	VIETNAM	88654	20.0	105.150803	9.173667	104.995228	8.757629

10421 rows x 11 columns

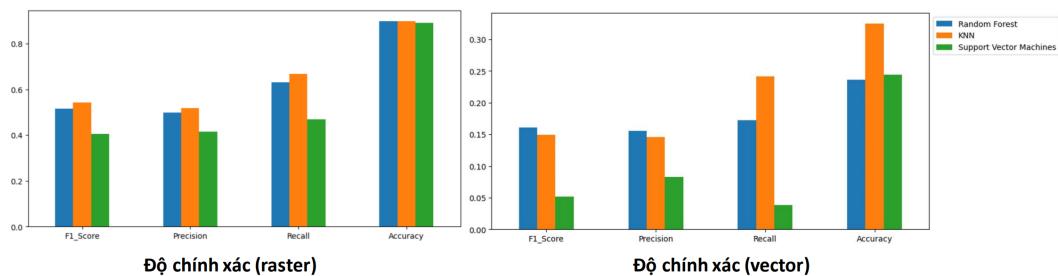
3.5 Phân tích kết quả



Bằng việc sử dụng dữ liệu raster trong việc dự đoán cho kết quả rất khả quan khi độ đo accuracy có thể đạt đến độ chính xác 90%

	precision	recall	f1-score	support
Tree cover	0.76	0.57	0.65	1749
Shrubland	0.54	0.51	0.52	467
Grassland	0.67	0.78	0.72	3205
Cropland	0.71	0.04	0.07	138
Built-up	0.78	0.88	0.83	4319
Bare/sparse vegetation	0.59	0.23	0.33	452
Permanent water bodies	0.99	0.99	0.99	18587
Mangroves	0.00	0.00	0.00	51
accuracy			0.90	28968
macro avg	0.63	0.50	0.51	28968
weighted avg	0.90	0.90	0.89	28968

Kết quả khi so sánh với dữ liệu vector



Qua việc phân tích so sánh giữa raster và vector, chúng ta thấy dữ liệu dạng raster sẽ cho kết quả tốt hơn khi xử lý bằng thị giác máy tính. Dữ liệu vector tuy có thể biểu diễn tốt hơn các khu vực, tuy nhiên, vì tính chất bộ dữ liệu khá hiếm và đặc thù khó biểu diễn một số tính chất theo yêu cầu bài toán nên kết quả khá thấp.

4 Thuật toán tính độ dốc nhanh bằng việc vector hóa và song song hóa cho kiến trúc đa lõi sử dụng dữ liệu DEM

4.1 Mức độ cần thiết

Ngày càng có nhiều nguyên cứu về land use và land cover. Trong đó các mô hình tính toán và các mô hình học máy phụ thuộc rất nhiều vào yếu tố độ dốc. Đặc biệt là các bài toán xử lý về tốc độ dòng chảy, độ sạt lở, thảm thấu đất... Tuy các bộ dữ liệu hiện tại thường chỉ cung cấp các thuộc tính về độ cao địa hình. Nếu muốn sử dụng thuộc tính độ dốc bắt buộc chúng ta phải tự tính toán thông qua các công cụ thông dụng hiện tại như qgis và agrgis. Nhưng đa phần các ứng dụng này lại đang

sử dụng phương pháp tính toán tuần tự cho raster gây mất nhiều thời gian. Việc đề xuất ra một phương pháp tính toán áp dụng tính song song hoá của vector để tăng tốc độ truy cập và thực thi dữ liệu là vô cùng cần thiết.

4.2 Công thức áp dụng

$$S = \arctan \sqrt{p^2 + q^2}$$

Để tính toán độ dốc p (theo hướng ngang - tây đông) và q (theo hướng dọc - nam bắc) cho mỗi ô trên bản đồ độ cao:

- Độ dốc p (theo hướng x) được tính bằng trung bình có trọng số của sai phân trung tâm theo hướng ngang của các ô lân cận.
- Độ dốc q (theo hướng y) được tính bằng trung bình có trọng số của sai phân trung tâm theo hướng dọc của các ô lân cận.

Các công thức này giúp xác định gradient của địa hình và thường được áp dụng trong các thuật toán xử lý bản đồ độ cao.

4.3 Chuyển đổi song song hoá

Thuật toán 1 BaseSlope (B): Thuật toán tuần tự cơ bản.

Đầu vào:

- dem — DEM đầu vào
- Δx — kích thước ô từ trái sang phải
- Δy — kích thước ô từ trên xuống dưới
- n — số hàng
- m — số cột

```

• for r ← 1... (n - 2) do
  •   for c ← 1... (m - 2) do
    •     p ← ((dem[r - 1][c + 1] + 2dem[r][c + 1]
      •       + dem[r + 1][c + 1])
    •     - (dem[r - 1][c - 1] + 2dem[r][c - 1]
      •       + dem[r + 1][c - 1]))
    •     / (8Δx)
    •     q ← (((dem[r + 1][c - 1] + 2dem[r + 1][c]
      •       + dem[r + 1][c + 1])
    •     - (dem[r - 1][c - 1] + 2dem[r - 1][c]
      •       + dem[r - 1][c + 1]))
    •     / (8Δy)
    •     slope[r][c] ← arctan( $\sqrt{p^2 + q^2}$ )
    •
  •   return slope

```

Đầu ra: slope - raster đầu ra có cùng kích thước

Thuật toán 2 TransformedSlope (T): Thuật toán tuần tự được biến đổi.

Đầu vào:

- dem — DEM đầu vào
- Δx — kích thước ô từ trái sang phải
- Δy — kích thước ô từ trên xuống dưới
- n — số hàng
- m — số cột

Output: slope - raster đầu ra có cùng kích thước

```

for r ← 1... (n - 2) do
    /* calculating p */
    for c ← 1... (m - 2) do
        p[c] ← dem[r - 1][c + 1] - dem[r - 1][c - 1]
    for c ← 1... (m - 2) do
        p[c] ← p[c]
        + 2(dem[r][c + 1] - dem[r][c - 1])
    for c ← 1... (m - 2) do
        p[c] ← (p[c] + (dem[r + 1][c + 1] - dem[r + 1][c - 1])) / (8Δx)

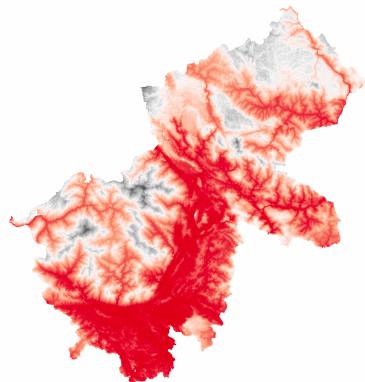
    /* calculating q */
    for c ← 1... (m - 2) do
        q[c] ← dem[r + 1][c - 1]
        + 2dem[r + 1][c] + dem[r + 1][c + 1]
    for c ← 1... (m - 2) do
        q[c] ← (q[c] - dem[r - 1][c - 1]
        + 2dem[r - 1][c]
        + dem[r - 1][c + 1]) / (8Δy)

    /* calculating slope */
    for c ← 1... (m - 2) do
        slope[r][c] ← arctan(√(p[c]^2 + q[c]^2))

return slope

```

4.4 Kết quả thực nghiệm và so sánh

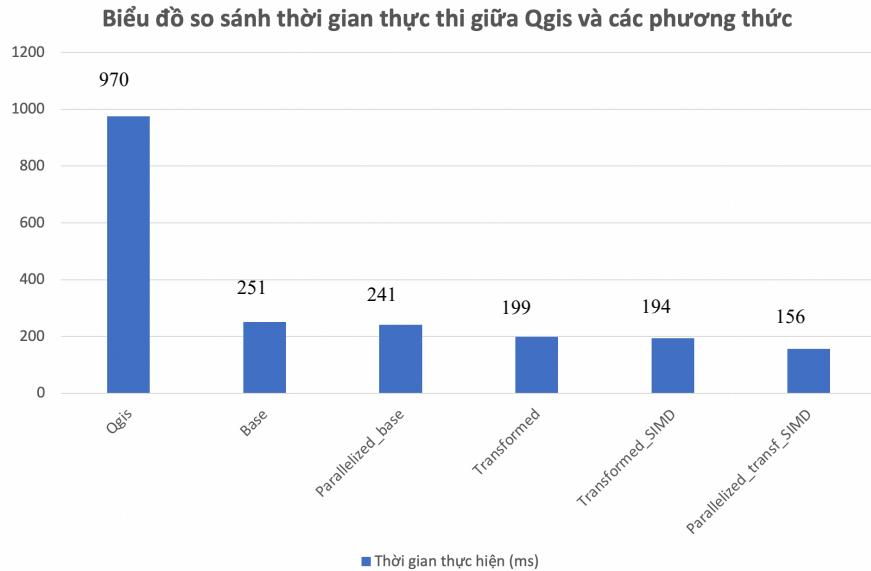


Hình 12 Dữ liệu DEM tỉnh Hà Giang

Bảng tốc độ thực thi của thuật toán tính độ dốc nhanh trên từng loại mô hình

	Reading time	Computing time	Writing time	Total time
Base	0.0686	0.1531	0.0291	0.251
Parallelized base	0.0923	0.1331	0.0165	0.241
Transformed	0.0692	0.1096	0.0201	0.199
Trans-formed with SIMD	0.0559	0.1034	0.0353	0.194
Parallelized_transformed_with SIMD	0.0359	0.1079	0.0131	0.156

Dữ liệu được tính toán tại tỉnh Hà Giang với phương pháp **Parallelized_transformed_with SIMD** là hiệu quả nhất với thời gian thực hiện chỉ 0.156s



So với các công cụ hỗ trợ thông dụng hiện tại đặc biệt là Qgis. Thời gian thực thi của Qgis là 0.97s chậm hơn so với phương pháp tối ưu nhất là Parallelized_transformed SIMD

5 Tổng kết và hướng phát triển

Từ các phân tích đánh giá phía trên chúng ta có thể thấy đối với dữ liệu vector tuy là dữ liệu lưu trữ vô cùng nhẹ và tốc độ tính toán nhanh hơn rất nhiều so với dữ liệu raster. Nhưng ở vector vẫn còn tỏ ra một số bất cập trong việc biểu diễn nhiều thuộc tính cùng một lúc.Thêm vào đó khi cho vào mô hình học máy thì áp dụng các kỹ thuật của raster thì kết quả sẽ được nâng cao hơn rất nhiều. Từ những ưu và nhược điểm trên, sẽ tiếp tục đề suất đưa ra hướng nghiên cứu mới về sự kết hợp của 2 dạng dữ liệu không chỉ trên việc tính toán độ dốc. Chúng tôi cũng kỳ vọng mô hình mới sở hữu được đặc trưng ưu việt của cả 2 dạng dữ liệu. Đặc biệt là trong khuôn khổ bài toán land use – land cover đang rất được chú ý hiện nay.

Tài liệu tham khảo

1. National Ocean Service, <https://oceanservice.noaa.gov/facts/lclu.html>, last accessed 2024/01/20.
2. Data by country, <https://www.diva-gis.org/gdata>, last accessed 2024/01/20.
3. Land-use and land-cover, <https://ec.europa.eu/eurostat/documents/3217494/5728541/KS-HA-11-001-12-EN.PDF.pdf/88b77a74-2628-4ed7-bb9f-24244d6c78a1?t=1414775978000>, last accessed 2024/01/20.
4. Eris land cover, <https://livingatlas.arcgis.com/landcover/>, last accessed 2024/01/20.
5. Land-use, land-cover change, https://archive.ipcc.ch/ipccreports/sres/land_use/index.php?idp=45, last accessed 2024/01/20.
6. Sự khác biệt giữa raster và vector, <https://idesign.vn/graphic-design/su-khac-biet-co-ban-giua-vector-va-raster-434345.html>, last accessed 2024/01/20.
7. Raster vs. Vector, <https://www.adobe.com/creativecloud/file-types/image/comparison/raster-vs-vector.html>, last accessed 2024/01/20.
8. Điểm khác biệt giữa ảnh vector và ảnh raster, <https://colorme.vn/blog/diem-khac-biet-giua-anh-vector-va-anh-raster>, last accessed 2024/01/20.
9. Prediction of Future Land Use/ Land Cover changes using a coupled CA-ANN model, <https://www.mdpi.com/2072-4292/15/4/1148>, last accessed 2024/01/20.
10. Land use and land cover change detection and prediction in Bhutan, <https://www.sciencedirect.com/science/article/pii/S2667010020300172>, last accessed 2024/01/20.
11. Land use and land cover prediction and its impact on surface runoff, <https://www.thaiscience.info/journals/Article/SJST/10973252.pdf>, last accessed 2024/01/20.