

Ứng dụng học máy để phân vùng ảnh và phát hiện bất thường trong giám sát rừng

Trần Văn Thành¹

GVHD: Phạm Văn Hải² Đỗ Phan Thuận¹

Hanoi University of Science and Technology



¹Computer Science

²Information System

January 16, 2019

Nội dung

- 1 Giới thiệu
- 2 Kiến thức cơ sở
- 3 Phương pháp đề xuất
- 4 Đánh giá
- 5 Kết luận

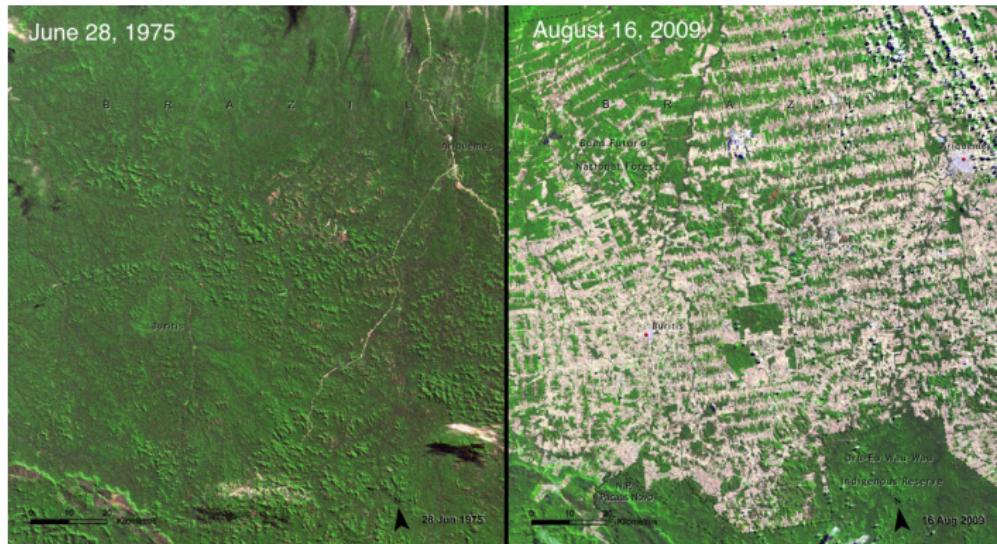
Giới thiệu

- Các vấn đề về biến đổi khí hậu

Giới thiệu

- Các vấn đề về biến đổi khí hậu
 - Quản lý tài nguyên, giám sát và bảo vệ rừng
 - Phát hiện những thay đổi bất thường

Giới thiệu



Rừng Rondonia, Brazil, tháng 6/1975 và tháng 8/2009.¹

¹ NASA

Giới thiệu



Hồ Mar Chiquita, Argentina, tháng 7/1998 và tháng 9/2011.²

² NASA

Giới thiệu

- Phương pháp truyền thống
 - Dùng bản vẽ, bản đồ
 - ĐI thực nghiệm

Giới thiệu

- Phương pháp truyền thống
 - Dùng bản vẽ, bản đồ
 - Đi thực nghiệm
- Hạn chế
 - Nguy hiểm, tốn thời gian và công sức
 - Không thể thực hiện đồng loạt trên quy mô lớn
 - Có những nơi con người không thể đến để kiểm tra đo đạc

Giới thiệu

- Phương pháp truyền thống
 - Dùng bản vẽ, bản đồ
 - Đi thực nghiệm
- Hạn chế
 - Nguy hiểm, tốn thời gian và công sức
 - Không thể thực hiện đồng loạt trên quy mô lớn
 - Có những nơi con người không thể đến để kiểm tra đo đạc

Bài toán đặt ra

- Phân vùng ảnh vệ tinh, phát hiện bất thường, trợ giúp ra quyết định trong giám sát và bảo vệ rừng

Công việc liên quan

Các phương pháp phân vùng ảnh

Công việc liên quan

Các phương pháp phân vùng ảnh

- Mô hình học sâu
 - ResNet, FCN, R-CNN,
U-Net,..
 - Hiệu quả, độ chính xác cao

Công việc liên quan

Các phương pháp phân vùng ảnh

- Mô hình học sâu
 - ResNet, FCN, R-CNN, U-Net,..
 - Hiệu quả, độ chính xác cao
- Phân cụm
 - K-means, Fuzzy C-Means, Mean-shift,..
 - Dễ hiểu, dễ cài đặt

Công việc liên quan

Các phương pháp phân vùng ảnh

- Mô hình học sâu
 - ResNet, FCN, R-CNN, U-Net,..
 - Hiệu quả, độ chính xác cao
- Phân cụm
 - K-means, Fuzzy C-Means, Mean-shift,..
 - Dễ hiểu, dễ cài đặt

Vấn đề

- Chỉ những thuật toán phức tạp mới có thể giải quyết được vấn đề !?

Công việc liên quan

- Những hạn chế của Deep Learning

Công việc liên quan

- **Những hạn chế của Deep Learning**

- i. Phụ thuộc vào số lượng và chất lượng dữ liệu
- ii. Cấu trúc mạng nơ-ron hoạt động phức tạp, khó hiểu
- iii. Tốn kém tài nguyên bộ nhớ, thời gian và chi phí

Công việc liên quan

- Những hạn chế của Deep Learning
 - ➊ Phụ thuộc vào số lượng và chất lượng dữ liệu
 - ➋ Cấu trúc mạng nơ-ron hoạt động phức tạp, khó hiểu
 - ➌ Tốn kém tài nguyên bộ nhớ, thời gian và chi phí
- Những thuật toán phức tạp
 - Yêu cầu độ tính toán cao
 - Nhạy cảm với cách chọn tham số đầu vào

Công việc liên quan

- Những hạn chế của Deep Learning
 - Phụ thuộc vào số lượng và chất lượng dữ liệu
 - Cấu trúc mạng nơ-ron hoạt động phức tạp, khó hiểu
 - Tốn kém tài nguyên bộ nhớ, thời gian và chi phí
- Những thuật toán phức tạp
 - Yêu cầu độ tính toán cao
 - Nhạy cảm với cách chọn tham số đầu vào
- Những thuật toán đơn giản
 - Sớm có một mô hình tổng quát cho bài toán
 - Kết quả sẽ được dần cải thiện ở những bước sau

Công việc liên quan

- Những hạn chế của Deep Learning
 - Phụ thuộc vào số lượng và chất lượng dữ liệu
 - Cấu trúc mạng nơ-ron hoạt động phức tạp, khó hiểu
 - Tốn kém tài nguyên bộ nhớ, thời gian và chi phí
- Những thuật toán phức tạp
 - Yêu cầu độ tính toán cao
 - Nhạy cảm với cách chọn tham số đầu vào
- Những thuật toán đơn giản
 - Sớm có một mô hình tổng quát cho bài toán
 - Kết quả sẽ được dần cải thiện ở những bước sau

Phương pháp học không giám sát

- Thuật toán phân cụm K-means

Algorithm 1 Thuật toán phân cụm K-means

1: **procedure** K-MEANS(D, k):

D - tập học

k - số cụm kết quả

2: Khởi tạo: Chọn ngẫu nhiên k quan sát trong tập D để làm các điểm trung tâm ban đầu

3: **while** not CONVERGENCE **do**

4: **for** each $x \in D$ **do**

5: Tính các khoảng cách từ x đến các điểm trung tâm ³

6: Gán x vào cụm có điểm trung tâm gần x nhất

7: **end for**

8: **for** each cụm **do**

9: Xác định lại điểm trung tâm dựa trên các quan sát hiện thời đang thuộc vào cụm này

10: **end for**

11: **end while**

12: **end procedure**

³ Khoảng cách Euclidean

Thuật toán phân cụm K-means

- Những hạn chế

Thuật toán phân cụm K-means

- Nhũng hạn chế
 - ① Cần biết trước số cụm k

Thuật toán phân cụm K-means

- Những hạn chế
 - 1. Cần biết trước số cụm k
 - 2. Kết quả và tốc độ hội tụ phụ thuộc vào bước khởi tạo

Thuật toán phân cụm K-means

- Nhũng hạn chế
 - 1. Cần biết trước số cụm k
 - 2. Kết quả và tốc độ hội tụ phụ thuộc vào bước khởi tạo
 - 3. Các cụm cần có dạng hình cầu, kích thước gần bằng nhau

Thuật toán phân cụm K-means

- **Những hạn chế**

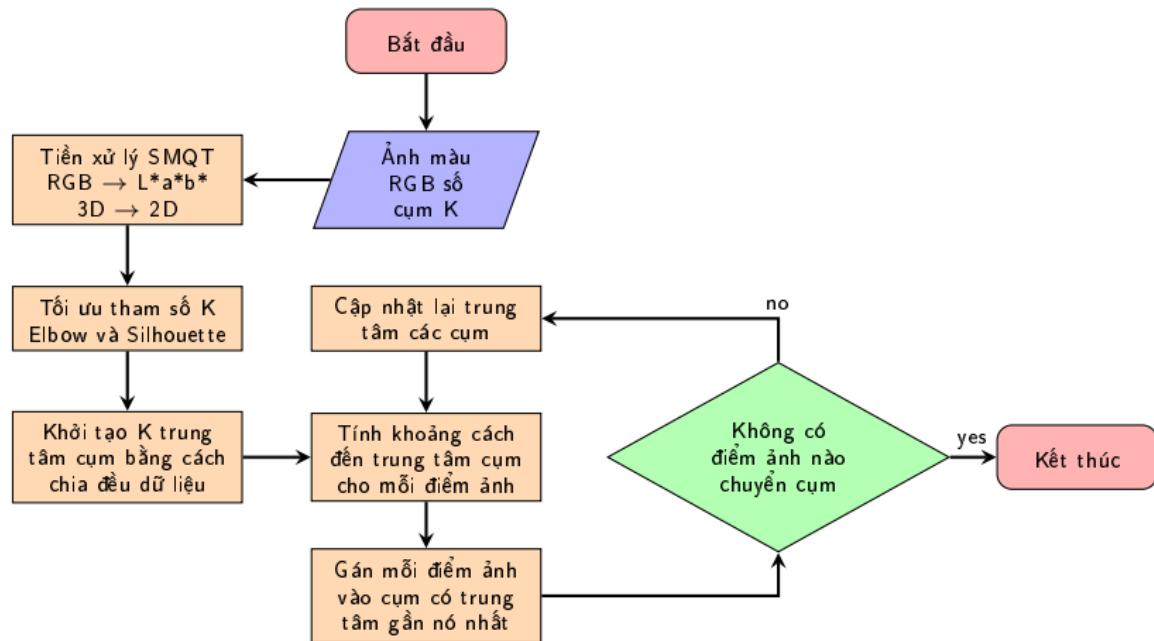
- i. Cần biết trước số cụm k
- ii. Kết quả và tốc độ hội tụ phụ thuộc vào bước khởi tạo
- iii. Các cụm cần có dạng hình cầu, kích thước gần bằng nhau
- iv. Không thể phân cụm dữ liệu nằm bên trong cụm khác

Các công việc cải tiến K-means trước đây

Một số đề xuất cải tiến K-means

Tác giả, năm	Phương pháp đề xuất
G. P. Babu, M. N. Murty, 1993	Lập trình di truyền dựa trên tâm cụm tối ưu gần nhất
C.Huang, R. Harris , 1993	Tìm kiếm trực tiếp bằng phương pháp chia nhị phân
I. Katsavounidis, C. C. J. Kuo, Z. Zhen, 1994	Khởi tạo tâm cụm ở các cạnh (biên) của dữ liệu
M. B. A. Daoud, S. A. Roberts, 1996	Khởi tạo dựa trên phân phối ngẫu nhiên trong cụm
A. Likas, N. Vlassis, J. J. Verbeek, 2003	Tăng dần số cụm cho đến khi tìm thấy K cho đến khi hội tụ
S. S. Khan, A. Ahmad, 2004	Khởi tạo tâm cụm dựa trên mật độ ngưng tụ dữ liệu đa quy mô
Bo Zhao and Zhongxiang Zhu, 2007	Phân vùng ảnh dựa trên tối ưu hóa đàm kiêng
Wang Min and Yin Siqing, 2010	Cải tiến K-means dựa trên giải thuật di truyền

Phương pháp đề xuất BK-means



Sơ đồ phương pháp BK-means.

Biểu diễn dữ liệu

			Array RGB								
			Page 3 -	blue	intensity	values	0.689	0.706	0.118	0.884	...
			green	intensity	values	0.535	0.532	0.653	0.925
			red	intensity	values	0.314	0.286	0.159	0.701
			Page 2 -	blue	intensity	values	0.553	0.633	0.528	0.493	...
			green	intensity	values	0.441	0.465	0.512	0.512
			red	intensity	values	0.342	0.647	0.515	0.816
			Page 1 -	blue	intensity	values	0.111	0.300	0.205	0.526	...
			green	intensity	values	0.523	0.428	0.712	0.926
			red	intensity	values	0.214	0.604	0.916	0.344
			Page 1 -	blue	intensity	values	0.100	0.121	0.173	0.126	...
							0.112	0.986	0.234	0.432	...
							0.785	0.128	0.863	0.521	...
							1.000	0.985	0.761	0.698	...
							0.455	0.783	0.224	0.395	...
							0.021	0.500	0.311	0.123	...
							1.000	1.000	0.867	0.051	...
							1.000	0.945	0.998	0.893	...
							0.990	0.941	1.000	0.876	...
							0.902	0.867	0.834	0.798	...
						
											...

Biểu diễn ảnh màu RGB bằng ma trận 3 chiều.⁴

Định dạng lại ảnh về ma trận 2 chiều:
 $\text{Image}(h,w,c) \mapsto \text{Matrix}(h \times w, c)$

⁴ Image 3D

Tiền xử lý ảnh

- Phương pháp SMQT⁵

- MQU⁶: Pixel (V) → Mean → M{0,1}

$$M(x) = \begin{cases} 1, & \text{if } V(x) > \text{Mean} \\ 0, & \text{else} \end{cases}$$

⁵ Successive Mean Quantization Transform

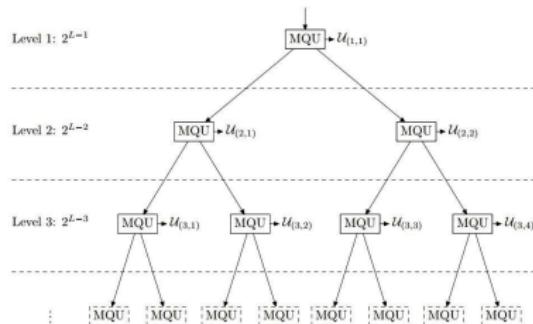
⁶ Mean Quantization Units

Tiền xử lý ảnh

- Phương pháp SMQT⁵

- MQU⁶: Pixel (V) → Mean → M{0,1}

$$M(x) = \begin{cases} 1, & \text{if } V(x) > \text{Mean} \\ 0, & \text{else} \end{cases}$$



Cây nhị phân SMQT.

⁵ Successive Mean Quantization Transform

⁶ Mean Quantization Units

Phương pháp SMQT

V		25		62		32		5		2		6		70		42		54		0		20		10	
---	--	----	--	----	--	----	--	---	--	---	--	---	--	----	--	----	--	----	--	---	--	----	--	----	--

Phương pháp SMQT

V	25	62	32	5	2	6	70	42	54	0	20	10
1	25	5	2	6	0	20	10	62	32	70	42	54
	0	0	0	0	0	0	0	1	1	1	1	1

Phương pháp SMQT

V	25	62	32	5	2	6	70	42	54	0	20	10
1	25 0	5 0	2 0	6 0	0 0	20 0	10 0	62 1	32 1	70 1	42 1	54 1
2	5 00	2 00	6 00	0 00	25 01	20 01	10 01	32 10	42 10	62 11	70 11	54 11

Phương pháp SMQT

V	25	62	32	5	2	6	70	42	54	0	20	10
1	25 0	5 0	2 0	6 0	0 0	20 0	10 0	62 1	32 1	70 1	42 1	54 1
2	5 00	2 00	6 00	0 00	25 01	20 01	10 01	32 10	42 10	62 11	70 11	54 11
3	2 000	0 000	5 001	6 001	10 010	25 011	20 011	32 100	42 101	54 110	62 110	70 111

Phương pháp SMQT

V	25	62	32	5	2	6	70	42	54	0	20	10
1	25 0	5 0	2 0	6 0	0 0	20 0	10 0	62 1	32 1	70 1	42 1	54 1
2	5 00	2 00	6 00	0 00	25 01	20 01	10 01	32 10	42 10	62 11	70 11	54 11
3	2 000	0 000	5 001	6 001	10 010	25 011	20 011	32 100	42 101	54 110	62 110	70 111
4	0 0000	2 0001	5 0010	6 0011	10 0100	20 0110	25 0111	32 1000	42 1010	54 1100	62 1101	70 1110

Phương pháp SMQT

V	25	62	32	5	2	6	70	42	54	0	20	10
1	25 0	5 0	2 0	6 0	0 0	20 0	10 0	62 1	32 1	70 1	42 1	54 1
2	5 00	2 00	6 00	0 00	25 01	20 01	10 01	32 10	42 10	62 11	70 11	54 11
3	2 000	0 000	5 001	6 001	10 010	25 011	20 011	32 100	42 101	54 110	62 110	70 111
4	0 0000	2 0001	5 0010	6 0011	10 0100	20 0110	25 0111	32 1000	42 1010	54 1100	62 1101	70 1110
5-8	0 0000	2 0001	5 0010	6 0011	10 0100	20 0110	25 0111	32 1000	42 1010	54 1100	62 1101	70 1110
	0000 0	0000 16	0000 32	0000 48	0000 64	0000 96	0000 122	0000 128	0000 160	0000 192	0000 208	0000 224

Tiền xử lý ảnh



Kết quả tiền xử lý ảnh bằng SMQT.

Lựa chọn tham số K

- Phương pháp khửu tay (Elbow)

- Tại vị trí khửu tay sẽ cho giá trị K tốt nhất
- Lỗi sẽ giảm rất nhanh (chậm) trước (sau) khi đi qua vị trí K

Lựa chọn tham số K

- Phương pháp khửu tay (Elbow)

- Tại vị trí khửu tay sẽ cho giá trị K tốt nhất
- Lối sẽ giảm rất nhanh (chậm) trước (sau) khi đi qua vị trí K

- Phương pháp hình chiêu (Silhouette)

$$a(o_i) = \frac{1}{|C_A| - 1} \sum_{o_j \in C_A, o_j \neq o_i} d(o_i, o_j)$$

$$b(o_i) = \min_{C_B \neq C_A} \frac{1}{|C_B|} \sum_{o_j \in C_B} d(o_i, o_j)$$

$$sil(o_i) = \frac{b(o_i) - a(o_i)}{\max\{b(o_i), a(o_i)\}}$$

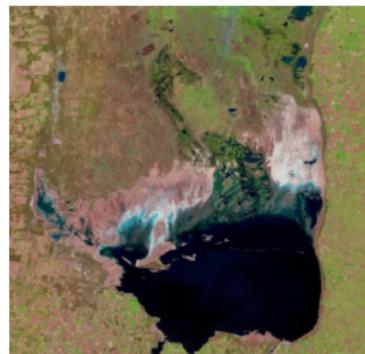
- Sự gắn kết của các điểm dữ liệu trong cùng cụm

- Sự tách biệt với các cụm lân cận

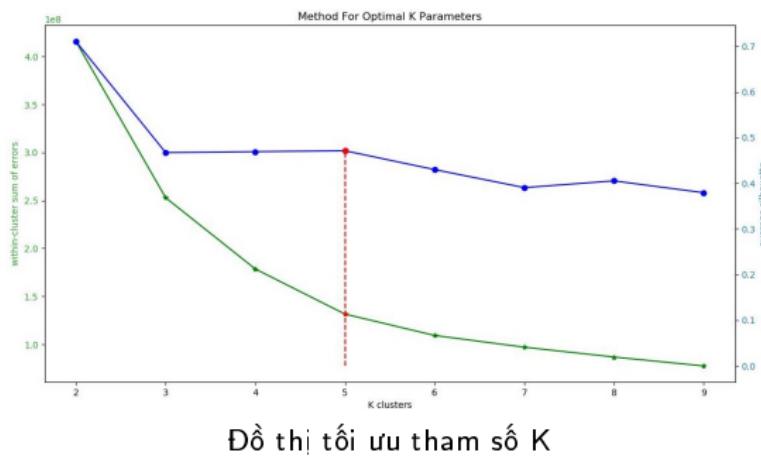
$$S_i = \begin{cases} 1 - a_i/b_i, & \text{if } a_i < b_i \\ 0, & \text{if } a_i = b_i \\ b_i/a_i - 1, & \text{if } a_i > b_i \end{cases}$$

$-1 \leq S_i \leq 1$

Lựa chọn tham số K



Ảnh gốc⁷ (493×481)



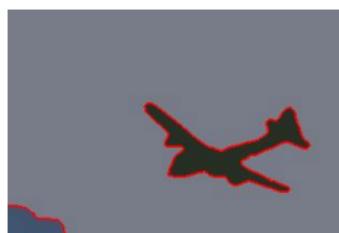
Kết quả tối ưu tham số K

K	2	3	4	5	6	7	8	9
$wcss \times 10^8$	4.1562	2.5291	1.7860	1.3163	1.0935	0.9713	0.8687	0.7756
silhouette	0.7101	0.4671	0.4691	0.4711	0.4298	0.3904	0.4051	0.3793

Thử nghiệm



Ảnh gốc⁸



Ground Truth



BK-means

Kết quả gán nhãn 481×321 pixel với $K=2$

Total=154401	Predicted as 0	Predicted as 1
Actual: 0	TP = 140693	FN = 609
Actual: 1	FP = 4145	TN = 8954

⁸ Berkeley

Thử nghiệm

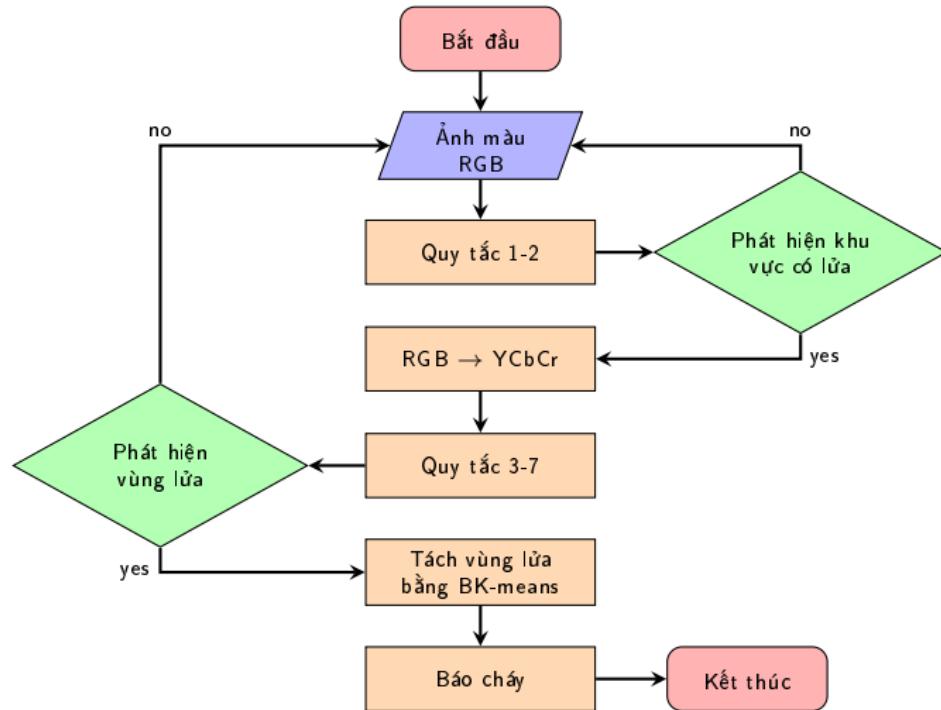
Kết quả phân lớp 481×321 pixel với K=2

	Same cluster	Different cluster
Same classes	TP = 9946045935	FN = 122796367
Different classes	FP = 588625471	TN = 1262289427

	Predicted as Positive	Predicted as Negative
Actual:Positive	TPR = 0.987804	FNR = 0.012196
Actual:Negative	FPR = 0.318019	TNR = 0.681981

RI	0.940316
Precision	0.944125
Recall	0.987804
F1	0.965471

Xây dựng mô hình



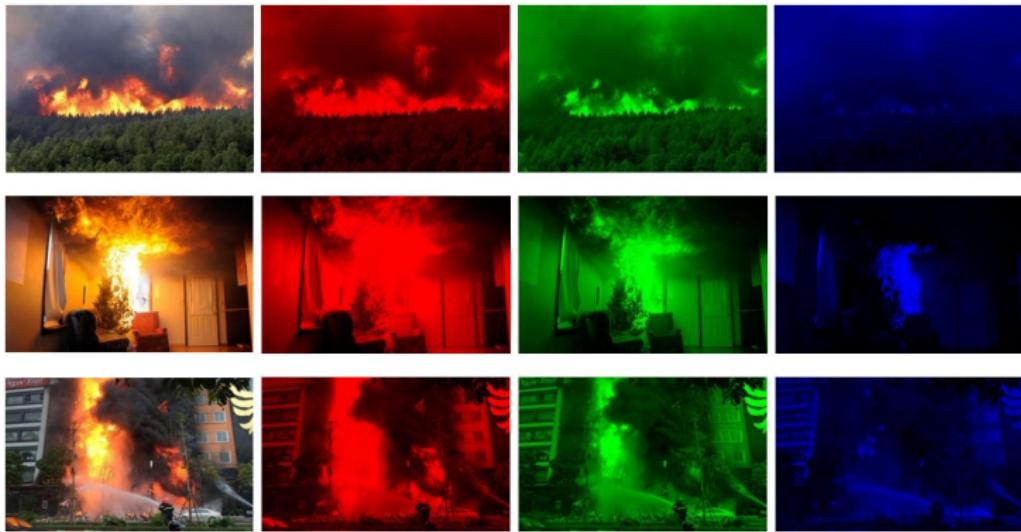
Sơ đồ phát hiện đám cháy trong ảnh.

Xây dựng mô hình

Xác định điểm ảnh vùng lửa

Hệ màu	Quy tắc
RGB	1. $R > G > B$ 2. $R > R_{mean} \cap G > G_{mean} \cap B > B_{mean}$
YCbCr	3. $Y \geq Cb$ 4. $Cr \geq Cb$ 5. $Y \geq Y_{mean} \cap Cb \geq Cb_{mean} \cap Cr \geq Cr_{mean}$ 6. $ Cb - Cr \geq Th, (Th=75)$ 7. $Cb \leq 150 \cap Cr \geq 120$

Xây dựng mô hình



Tách các kênh màu R,G,B.⁹

⁹ Quy tắc 1,2

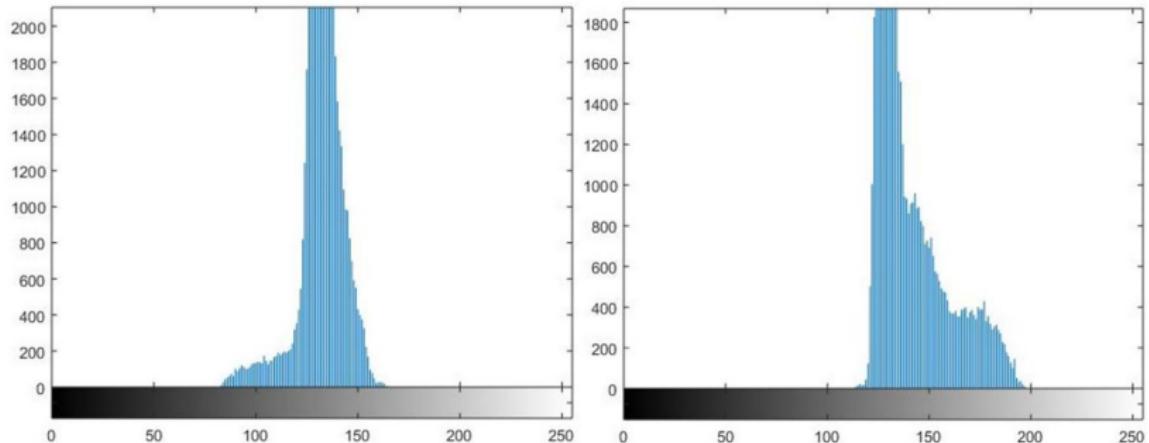
Xây dựng mô hình



Tách kênh sáng Y, kênh màu Cb và Cr.¹⁰

¹⁰ Quy tắc 3,4,5

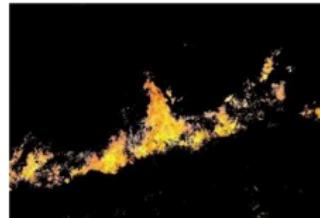
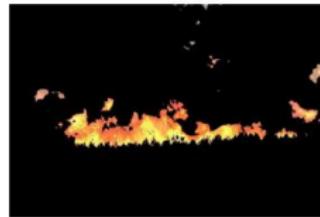
Xây dựng mô hình



Biểu đồ màu trên kênh Cb và Cr.¹¹

¹¹Quy tắc 7

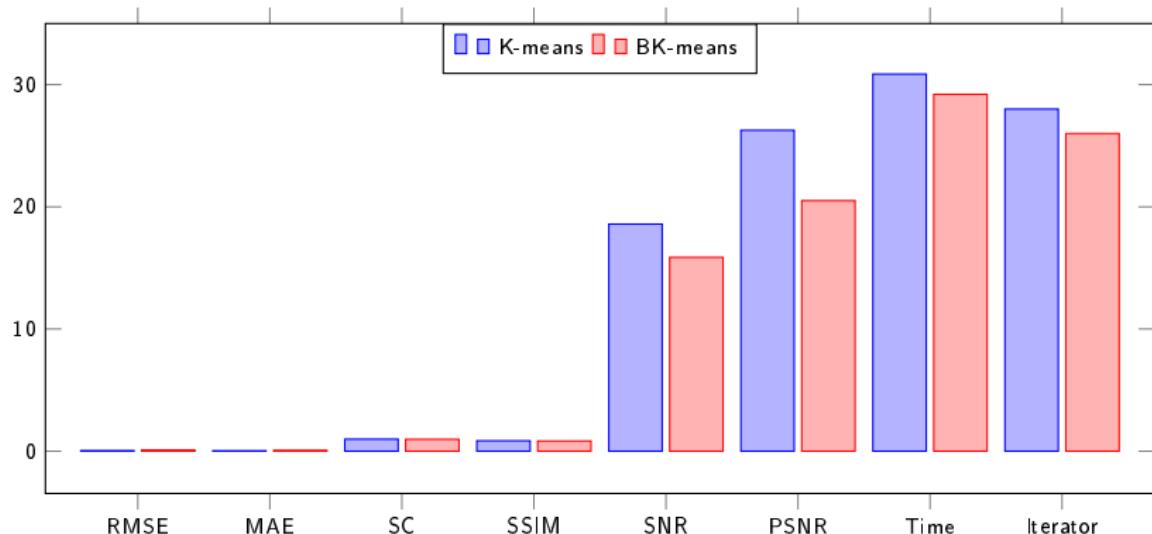
Thử nghiệm



Thử nghiệm nhận diện đám cháy trong ảnh.

BK-means

	RMSE	MAE	SC	SSIM	SNR	PSNR	Time	Iterator
K-means	0.0486	0.0362	0.9863	0.8550	18.5860	26.2638	30.8658	28
BK-means	0.0944	0.0735	0.9768	0.8369	15.8679	20.5045	29.2071	26

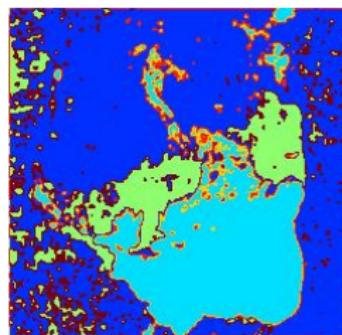


So sánh hiệu năng K-means và BK-means.

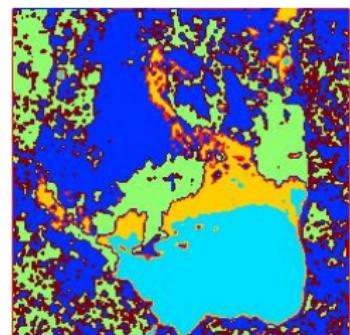
BK-means



Ảnh gốc



K-means

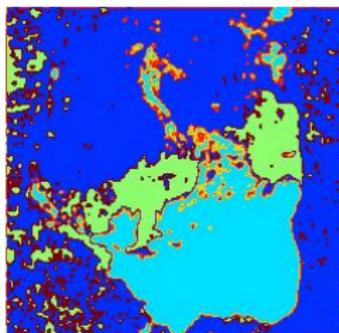


BK-means

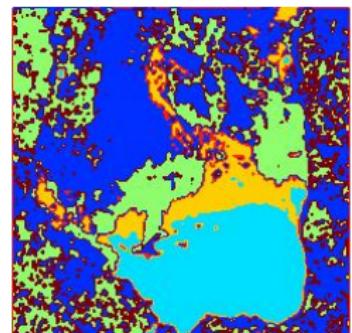
BK-means



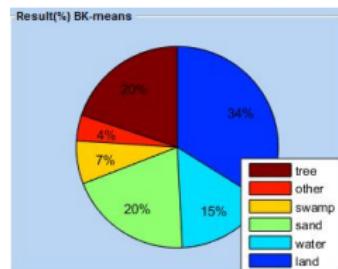
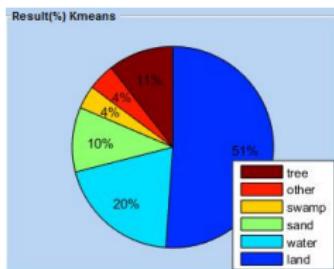
Ảnh gốc



K-means



BK-means



Phân loại ảnh

- Tập dữ liệu gồm 2 lớp:
 - fire (2400 ảnh)
 - not-fire (2931 ảnh)

Phân loại ảnh

- Tập dữ liệu gồm 2 lớp:
 - fire (2400 ảnh)
 - not-fire (2931 ảnh)

Kết quả phân loại ảnh

Total=5331	Predicted as Positive(fire)	Predicted as Negative(not-fire)
Actual:Positive(fire)	TP = 1689	FN = 711
Actual:Negative(not-fire)	FP = 566	TN = 2365
Actual:Positive(fire)	TPR = 0.7038	FNR = 0.2962
Actual:Negative(not-fire)	FPR = 0.1931	TNR = 0.8069
RI	0.7633	
Precision	0.7490	
Recall	0.7038	
F1	0.7257	

Kết quả đạt được

- Cải thiện của BK-means so với K-means
 - Thời gian chạy nhanh hơn
 - Số lần lặp rút ngắn
- Tỷ lệ 70.38% báo cháy đúng và 19.31% báo cháy giả

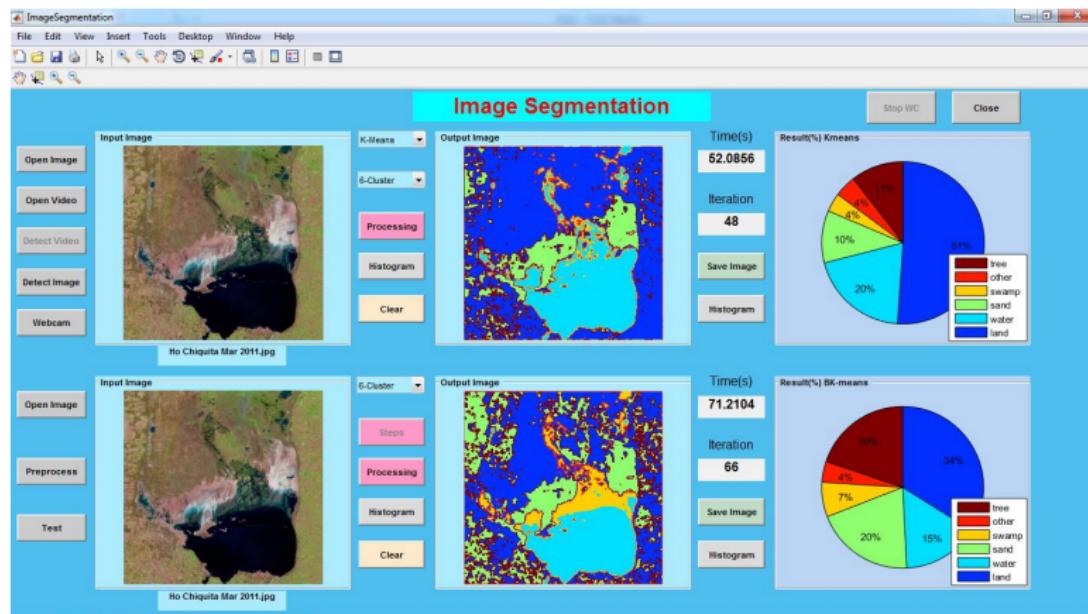
Kết quả đạt được

- Cải thiện của BK-means so với K-means
 - Thời gian chạy nhanh hơn
 - Số lần lặp rút ngắn
- Tỷ lệ 70.38% báo cháy đúng và 19.31% báo cháy giả
- Những hạn chế
 - Tập dữ liệu còn đơn giản, chưa đủ lớn
 - Nhận diện đám cháy trong video realtime còn chậm
 - Chưa nhận diện được khói

Kết quả đạt được

- Cải thiện của BK-means so với K-means
 - Thời gian chạy nhanh hơn
 - Số lần lặp rút ngắn
- Tỷ lệ 70.38% báo cháy đúng và 19.31% báo cháy giả
- Những hạn chế
 - Tập dữ liệu còn đơn giản, chưa đủ lớn
 - Nhận diện đám cháy trong video realtime còn chậm
 - Chưa nhận diện được khói
- Công việc tương lai
 - Phát hiện chuyển động của lửa và khói trong video để tăng tỷ lệ báo cháy đúng và giảm tỷ lệ báo cháy giả

Kết quả đạt được



Giao diện chương trình.

Q&A

Thank You!

Tài liệu tham khảo

- Vũ Hữu Tiệp, Machine Learning cơ bản, Nhà xuất bản khoa học và kỹ thuật, 2018.
- Zhongwu Wang, John R. Jensen, Jungholm., An automatic region-based image segmentation algorithm for remote sensing applications, Environmental Modelling & Software, Elsevier Science Publishers, Volume 25, Issue 10, October 2010, Pages 1149-1165.
- ZhenSong Chen, ZhiQuan Qi, Fan Meng, Limeng Cui & Yong Shi., Image Segmentation via Improving Clustering Algorithms with Density and Distance, Information Technology and Quantitative Management (ITQM 2015), Elsevier Science Publishers, 55: 1015-1022, 2015.
- V.V.