

2025-04-23 22:04

Tags:

Early Wildfire Detection with CubeSat Images Using Single Image Super-Resolution

- **Abstract summary:** Bài báo trình bày một phương pháp mới để phát hiện cháy rừng sớm bằng hình ảnh từ vệ tinh CubeSat, kết hợp giữa **deep learning** và **kỹ thuật tăng độ phân giải (super-resolution)**.
 - Note: **Dữ liệu là ảnh từ vệ tinh Landsat-8**, được xử lý thành ảnh RGB và **giả lập ảnh CubeSat** bằng cách **giảm chất lượng**, rồi **tăng lại độ phân giải bằng Real-ESRGAN**. **Mục tiêu** là tạo ra một giải pháp có thể áp dụng tốt **cho ảnh CubeSat trong thực tế**
- **Cách tiếp cận:**
 - Dùng ảnh RGB được tạo từ ảnh Landsat-8 (ban đầu có 10 kênh phổ, sau đó chuyển thành 3 kênh RGB).
 - Ảnh được nâng cấp độ phân giải gấp 4 lần bằng kỹ thuật **Real-ESRGAN**.
 - Áp dụng transfer learning với hai mô hình pre-trained: **MobileNetV2** và **ResNet152V2**.
- **Kết quả:** Việc nâng cấp độ phân giải ảnh giúp tăng độ chính xác, độ nhạy (recall) và f1-score khi phát hiện cháy rừng, tăng khoảng **3-5%** tùy mô hình. **### Introduction Summary**
- Nhắc lại về các tác hại của cháy rừng và tính cấp thiết của việc phát triển mô hình phát hiện cháy rừng
- Trình bày lại các phương pháp phát hiện cháy rừng trong quá khứ, có 3 loại:
 - dựa trên mặt đất (*terrestrial-based*), trên không (*aerial-based*) và vệ tinh (*satellite-based*)
 - Dạng mặt đất và trên không phổ biến hơn vì chi phí ban đầu và kỹ thuật đơn giản hơn.
 - Tuy nhiên, do số lượng vệ tinh phóng lên tăng mạnh và chi phí giảm, nghiên cứu về vệ tinh đang được đẩy mạnh.
- Trình bày về các ưu điểm của việc sử dụng vệ tinh:
 - Có thể giám sát những khu vực xa xôi, khó tiếp cận.
 - Quan sát liên tục cả vào ban đêm hoặc trong thời tiết xấu.
 - Về lâu dài tiết kiệm chi phí bảo trì so với các phương pháp khác.
- Trình bày cụ thể về những ưu điểm của **CubeSat**:
 - Giá rẻ, dễ triển khai, mở rộng vùng quan sát trên toàn cầu.
 - Có thể chọn quỹ đạo để theo dõi nhiều nơi hoặc giám sát

- liên tục một khu vực.
- Có thể dùng nhiều CubeSat để vượt qua hạn chế về băng thông truyền tín hiệu.
- Những hạn chế của CubeSat:
 - Kích thước nhỏ nên camera payload capacity nhỏ → không thể sử dụng thuật toán so sánh nhiều băng tần như các vệ tinh lớn.
 - Khả năng xử lý phần mềm giới hạn → không thể implementing các mô hình deep learning nặng.
- Do đó, cần giải pháp nhẹ, hiệu quả và dùng ít dữ liệu đầu vào.
 - => Bài báo đề xuất dùng kỹ thuật **super-resolution** để cải thiện hiệu quả phát hiện cháy rừng từ ảnh CubeSat, dù bị giới hạn về số băng tần và bộ nhớ. ### Super-resolution technique:
- Kỹ thuật super-resolution sẽ được thực hiện tại trạm ở mặt đất, mà không cần phải thay đổi bất kỳ phần cứng nào của CubeSat
- Mục tiêu là giúp hệ thống **phát hiện cháy rừng thời gian thực trên toàn cầu bằng CubeSat**.
- Ưu điểm:
 - Khắc phục các giới hạn của CubeSat (kích thước nhỏ, băng thông thấp, tải trọng hạn chế).
 - Ảnh RGB từ CubeSat được xử lý nâng cao độ phân giải ở mặt đất → cải thiện chất lượng ảnh.
 - Khi so sánh mô hình học sâu trên ảnh gốc và ảnh đã enhanced, kết quả cho thấy:
 - * **Tốc độ học nhanh hơn**
 - * **Hiệu suất phát hiện cháy tốt hơn** ### Materials
- Dữ liệu dùng cho training được tiền xử lý từ data Landsat-8
 - Ảnh Landsat-8 có 11 băng tần (multispectral), nhưng loại bỏ band 8 (panchromatic), còn lại 10 band → lưu dưới định dạng TIFF.
- Trong nghiên cứu trước đó:
 - Ảnh được *resize* về kích thước 256x256, với spatial resolution là 30m thì 1 bức ảnh sẽ tương ứng với $59km^2$ diện tích mặt đất [[Pasted image 20250424124416.png]]
 - Dùng kỹ thuật segmentation để tạo **fire masks** cho ảnh (xem rằng liệu có đang xảy ra 1 vụ cháy nào trong từng pixel không)
 - Tuy nhiên, gán nhãn thủ công rất khó và tốn thời gian → họ dùng **3 thuật toán phát hiện cháy có sẵn** để tự động tạo nhãn:
 - * Schroeder et al.
 - * Murphy et al.

* Kumar & Roy

- “Since the three algorithms are not ground truth, they sometimes produce slightly different results”, nên họ **xác định một pixel là “cháy” nếu ít nhất 2 thuật toán đồng ý.** ### Preprocessing: [[Pasted image 20250424144817.png]]
 - Do hạn chế về khả năng xử lý của CubeSat, kích thước ảnh được giảm xuống 64x64 pixel và chỉ sử dụng 3 kênh màu RGB thay vì 10 kênh đa phổ. Bài toán được chuyển từ phân đoạn từng pixel sang phân loại nhị phân (có cháy/không cháy).
1. Giảm kích thước ảnh:
 - Ảnh đã được thu nhỏ xuống kích cỡ 64x64 để phù hợp với khả năng xử lý của CubeSat, mỗi ảnh giờ đây presents $3.7km^2$
 2. Đổi format:
 - Từ ảnh multispectral 10 band TIFF format -> ảnh 3 band RGB với PNG format
 - Sử dụng thư viện GDAL của python để process
 - Mặc dù có mất một số thông tin khi chuyển từ giá trị float sang integer ở từng pixel, nhưng ảnh hưởng không đáng kể vì mô hình sử dụng thông tin đã được chuẩn hóa
 3. Chia nhỏ và gán nhãn lại:
 - a. Có ảnh gốc 256x256 và một fire mask tương ứng
 - Trong fire mask: Màu trắng = có cháy, màu đen = không cháy
 - Họ chia ảnh gốc thành 16 ảnh nhỏ (64x64) và cũng chia fire mask tương ứng thành 16 phần
 - b. Cách gán nhãn mới:
 - Nếu phần fire mask của ảnh nhỏ có bất kỳ điểm trắng nào → ảnh đó được gán nhãn “có cháy”
 - Nếu phần fire mask hoàn toàn đen → ảnh đó được gán nhãn “không cháy”
 - Assume rằng độ chính xác của các mặt nạ cháy gốc là ground truth và sử dụng chúng làm cơ sở để tạo bộ dữ liệu phân loại nhị phân mới (chỉ có 2 lớp: có cháy hoặc không có cháy). Điều này đơn giản hóa bài toán để phù hợp với khả năng xử lý hạn chế của CubeSat. ### Class imbalance
- Dữ liệu gốc được thiết kế cho pixel-level segmentation nên ngay cả ảnh chỉ có 1 pixel được classified là cháy thì cũng sẽ được labeled “có cháy”
 - Điều này tạo ra class imbalance nghiêm trọng với tỉ lệ 10:1
 - khi data imbalanced, mô hình có xu hướng biased theo majority class bởi vì nó đang cố tối ưu hóa độ chính xác tổng thể, dẫn đến hiệu suất kém trong việc phát hiện cháy rừng (điều quan trọng

nhất)

- **Giải Pháp:** sử dụng undersampling, giảm số lượng mẫu từ lớp đa số (giảm số ảnh “không cháy”) với các ảnh được chọn ngẫu nhiên để tránh mất thông tin quan trọng
- **Kết quả:**
 - Bộ dữ liệu cuối cùng dành cho binary classification deep-learning có 5.966 ảnh với tỷ lệ cân bằng giữa ảnh “có cháy” và “không cháy”
 - Cùng một tập hình ảnh được sử dụng cho cả quá trình huấn luyện và kiểm tra, không phụ thuộc vào việc có sử dụng kỹ thuật siêu phân giải hay không ### Methods: [[Pasted image 20250424144944.png]]
- **Early Wildfire Detection Framework**
 1. *Vấn đề với CubeSat:* Do CubeSat có hạn chế về khả năng xử lý bởi vì payload capacity nhỏ, điều này gây khó khăn cho việc xử lý và detect wildfire onboard
 - => Giải pháp tốt hơn được đề xuất là gửi ảnh về mặt đất để xử lý
 2. *Vấn đề mới:* Ngay cả khi xử lý trên mặt đất, ảnh từ CubeSat có độ phân giải cố định (fixed resolution). Chỉ dùng các mô hình phức tạp hơn không đủ để cải thiện hiệu suất vì chất lượng ảnh đầu vào bị giới hạn.
 - => Đề xuất *super-resolution technique*, giúp tăng cường chất lượng hình ảnh, làm ảnh nét hơn, chi tiết hơn từ ảnh gốc có độ phân giải thấp.
 - Nghiên cứu sẽ so sánh hiệu quả phát hiện cháy rừng khi dùng SR so với khi không dùng SR (trên cùng ảnh và mô hình, flow xử lý dựa theo Figure2)
 3. Khó khăn khi áp dụng SR cho ảnh vệ tinh:
 - Hầu hết nghiên cứu SR phát triển trong lĩnh vực thị giác máy tính, tập trung vào ảnh RGB.
 - Ảnh vệ tinh thường có nhiều kênh phổ multi-spectrum, bao gồm cả kênh hồng ngoại (infrared) rất hữu ích cho việc phát hiện cháy rừng. Do đó, các mô hình SR được pre-trained thường không phù hợp.
 - *Tuy nhiên*, nghiên cứu này vẫn áp dụng được SR vì đã có bước preprocess ảnh vệ tinh để chuyển đổi thành ảnh RGB
- **Giải thích về việc chọn Single-image super-resolution over multi-image super-resolution**
 - Do **băng thông vệ tinh hạn chế** và việc **thu thập ảnh vệ tinh chụp liên tiếp cùng một đám cháy là cực kỳ khó khăn**, nghiên cứu đã chọn phương pháp **Siêu phân giải Đơn ảnh (SISR)** thay vì Đa ảnh (MISR).
 - Họ sử dụng mô hình học sâu pre-trained **Real-ESRGAN**

để **tăng độ phân giải ảnh lên 4 lần**. Kết quả là chất lượng ảnh vệ tinh được cải thiện đáng kể, với chi tiết tốt hơn (độ phân giải hiệu quả tăng từ 30m lên 7.5m), giúp ích cho việc phát hiện cháy rừng. [[Pasted image 20250424163256.png]]

- **Wildfire Detection Learning Models:**

- Sử dụng 2 mô hình học sâu là: **MobileNetV2** và **ResNet152V2**.
- **Số lượng tham số:** MobileNetV2 có khoảng 2.4 triệu tham số, trong khi ResNet152V2 có tới 58.6 triệu.
- **Kích thước Mô hình (Model Size - MB):** Kích thước lưu trữ của MobileNetV2 chỉ là 9.75 MB, nhỏ hơn đáng kể so với 234.44 MB của ResNet152V2.

- **Training và thiết lập tham số mô hình:**

- Tiến hành **hai lượt train và test riêng biệt** cho *mỗi* mô hình (tức là tổng cộng 4 lần huấn luyện chính).
- Quá trình huấn luyện kéo dài tối đa **100 epochs**, ảnh đầu vào được qua các bước **normalizing, data augmentation** (chỉ lật ảnh, không cắt/xoay), và chọn mô hình tốt nhất dựa trên **validation loss**. Các siêu tham số cụ thể như trình tối ưu hóa Adam, tốc độ học $1e-5$, và **hàm loss Binary Cross-entropy** đã được chọn sau khi thử nghiệm. [[Pasted image 20250424195347.png]]

Result

- **Evaluation Metrics:**

- Sử dụng các metrics chính là: **Precision, Recall và F1-score**
- Sử dụng **ROC Curve và chỉ số AUC** để biểu diễn hiệu suất mô hình [[Pasted image 20250424201112.png]]

- **Performance on trained model**

- 4 combinations của 2 mô hình (*MobileNetV2* và *ResNet152V2*) cùng với 2 kiểu data (*Low-res* và *Super-res*) được đánh giá trên một bộ dữ liệu kiểm tra riêng biệt gồm 1,194 ảnh/bộ dữ liệu.
- Kết quả: [[Pasted image 20250424201940.png]] => Các mô hình sử dụng ảnh đã được *thực hiện super-resolution* cho thấy hiệu suất *vượt trội* so với các mô hình sử dụng ảnh gốc (độ phân giải thấp - *Low resolution*).
- Learning performance trên các model Low-res tăng khi model size tăng, ngược lại learning performance trên các model super-res không thay đổi khi model size tăng => Điều này cho thấy ảnh super-res chứa *nhiều thông tin rõ ràng hơn*. Ngay cả mô hình nhỏ hơn (*MobileNetV2*) cũng có thể học được gần như tất cả các đặc trưng cần thiết

từ super-res training data này, không thua kém nhiều so với mô hình lớn. => Khi chất lượng ảnh đầu vào đã tốt (super-res), việc *tăng kích thước training data* có thể quan trọng hơn việc *chỉ tăng model size*.

[[Pasted image 20250424202702.png]]

- Phân tích sâu hơn cho thấy các mô hình có xu hướng báo cháy nhầm (FP) nhiều hơn bỏ sót (FN), và điều này có thể được khắc phục bằng cách điều chỉnh ngưỡng hoặc tăng dữ liệu. Mặc dù P/R/F1 tương đương, điểm AUC cao hơn của ResNet cho thấy nó có khả năng phân biệt tổng thể tốt hơn, và có thể đạt hiệu suất cao hơn nếu tối ưu ngưỡng quyết định. ### Conclusion
Summary
- Bài báo kết luận rằng việc kết hợp *Deep learning* và *Super-res* là một hướng đi *hiệu quả* để phát hiện cháy rừng sớm bằng ảnh thu được từ CubeSat, khắc phục các hạn chế của các budget satellites. Phương pháp này sử dụng *transfer learning* bằng cách tận dụng các mô hình như MobileNetV2/ResNet152V2 và ảnh được tăng cường chất lượng nhờ *Super-res*, đã cho thấy *cải thiện rõ rệt (3-5%)* về hiệu suất so với dùng ảnh gốc. Nó mở ra tiềm năng cho hệ thống giám sát cháy toàn cầu, thời gian thực, chi phí thấp. Các hướng phát triển tiếp theo bao gồm *mở rộng bộ dữ liệu, thử nghiệm các hệ số Super-res khác nhau, tối ưu hóa ngưỡng quyết định cháy*, và đặc biệt là *giảm kích thước mô hình cùng tỷ lệ bỏ sót đám cháy* để phù hợp với việc xử lý trên CubeSat. # References