



Trích xuất cây xanh đô thị từ ảnh UAV bằng kỹ thuật học sâu phục vụ công tác quản lý đô thị

Trần Ngọc Huyền Trang^{1,2*}, Chung Phước Sang¹, Trần Khánh Luân¹

¹Khoa Trắc địa, Bản đồ và Công trình, Đại học Tài nguyên và Môi trường TP. HCM

²Khoa Môi trường và Tài nguyên, Đại học Bách Khoa TP. HCM

Email tác giả liên hệ: tnhtrang@hcmunre.edu.vn

doi.org/10.5281/zenodo.17068801

Tóm tắt:

Bài báo đề xuất một phương pháp ứng dụng kỹ thuật học sâu nhằm trích xuất cây xanh đô thị từ ảnh UAV có độ phân giải cao. UAV đóng vai trò thiết yếu trong việc cung cấp ảnh trực giao và ảnh xiên với độ chi tiết hình học cao, hỗ trợ mô hình nhận dạng hiệu quả không gian ba chiều của cây xanh. Trọng tâm của phương pháp là khả năng tự động nhận dạng và phân vùng cây xanh trên các đặc trưng hình ảnh học được, đặc biệt là hình dạng và kích thước tán cây, vốn có biên không đều và phân bố tự nhiên, khác biệt rõ rệt so với các đối tượng nhân tạo. Kết quả thu được là lớp dữ liệu không gian phản ánh chính xác vị trí và cấu trúc hình học của cây xanh, phục vụ hiệu quả cho công tác quản lý, phân tích và quy hoạch không gian xanh đô thị. Phương pháp cho thấy tính khả thi cao, khả năng tự động hóa và mở rộng tốt, góp phần xây dựng hệ thống quản lý đô thị thông minh và bền vững.

Từ khóa: UAV, Deep Learning, GIS, Cây xanh đô thị

Ngày nhận bài: 18/08/2025 Ngày sửa lại: 20/08/2025 Ngày chấp nhận đăng: 25/08/2025 Ngày xuất bản: 30/08/2025

Urban tree extraction from UAV Imagery using deep learning techniques for urban management applications

Trần Ngọc Huyền Trang^{1,2*}, Chung Phước Sang¹, Trần Khánh Luân¹

¹Faculty of Surveying, Mapping and Construction, Ho Chi Minh City University of Natural Resources and Environment

²Faculty of Environment and Resources, Ho Chi Minh City University of Technology (HCMUT)

Corresponding Author Email: tnhtrang@hcmunre.edu.vn

Abstract:

This paper proposes a method that leverages deep learning techniques to extract urban trees from high-resolution UAV imagery. UAVs play a crucial role in capturing both nadir and oblique images with high geometric detail, facilitating effective modeling of the three-dimensional spatial characteristics of urban vegetation. The core of the proposed approach lies in its ability to automatically identify and segment trees based on learned visual features, particularly the shape and size of tree canopies, which often exhibit irregular boundaries and natural spatial distribution, distinct from man-made objects. The resulting output is a spatial dataset that accurately reflects the location and geometric structure of trees, effectively supporting urban green space management, analysis, and planning. The method demonstrates high feasibility, automation capability, and scalability, contributing to the development of intelligent and sustainable urban management systems.

Keywords: UAV, Deep Learning, GIS, Urban tree

Submission received: 18/08/2025 Revised: 20/08/2025 Accepted: 25/08/2025 Published: 30/08/2025

1. Giới thiệu

Cây xanh đô thị đóng vai trò quan trọng trong việc cải thiện chất lượng không khí, giảm hiệu ứng đảo nhiệt, và nâng cao chất lượng sống cho cư dân đô thị [1]. Để quản lý hiệu quả cây xanh trong bối cảnh đô thị hóa nhanh, cần áp dụng các phương pháp hiện đại có khả năng tự động hóa cao và độ chính xác lớn trong giám sát, kiểm kê và quy hoạch.

Trên thế giới, nhiều quốc gia đã tích cực ứng dụng các công nghệ như hệ thống thông tin địa lý (GIS), viễn thám và máy bay không người lái (UAV) trong công tác quản lý cây xanh. Tại Hoa Kỳ, dự án Urban Tree Canopy Assessment sử dụng ảnh vệ tinh và dữ liệu LiDAR nhằm đánh giá độ che phủ tán cây ở cấp đô thị [1]. Singapore, thông qua Cơ quan Công viên Quốc gia (NParks), đã triển khai hệ thống quản lý cây xanh từ xa,



tích hợp công nghệ GIS và cảm biến để nâng cao độ chính xác và hiệu quả trong công tác kiểm kê, giám sát [2]. Các giải pháp công nghệ này không chỉ giúp nâng cao hiệu quả quản lý mà còn góp phần giảm chi phí vận hành [3].

Tại Việt Nam, quá trình quản lý cây xanh đô thị vẫn còn nhiều thách thức, đặc biệt trong bối cảnh đô thị hóa diễn ra nhanh chóng. Riêng TP. Hồ Chí Minh hiện có trên 200000 cây xanh phân bố dọc theo khoảng 1200 tuyến đường [4, 5], nhưng việc kiểm kê và quản lý phần lớn vẫn dựa vào phương pháp thủ công, thiếu tính đồng bộ và hiệu quả. Một số nghiên cứu bước đầu đã ứng dụng công nghệ UAV kết hợp với hệ thống GIS để xây dựng cơ sở dữ liệu cây xanh và hỗ trợ quản lý không gian xanh đô thị [6, 7]. Những kết quả ban đầu cho thấy tiềm năng rõ rệt trong việc nâng cao hiệu quả giám sát và quy hoạch, song việc ứng dụng các kỹ thuật học sâu (Deep Learning) trong trích xuất cây xanh từ ảnh UAV vẫn còn hạn chế, đặc biệt tại những khu vực đô thị phức tạp như bờ kênh Tẻ, Quận 7 (cũ) – nơi cây xanh phân bố xen kẽ với công trình kiến trúc và mặt nước.

Trong bối cảnh đó, kỹ thuật học sâu, đặc biệt là mạng nơ-ron tích chập (Convolutional Neural Network – CNN), đã chứng minh khả năng vượt trội trong phân tích ảnh, hỗ trợ hiệu quả các tác vụ như phân đoạn và phát hiện đối tượng [8]. Khi kết hợp với dữ liệu ảnh UAV có độ phân giải cao, các mô hình học sâu có thể tự động nhận diện cây xanh dựa trên hình dạng tán lá, màu sắc và vị trí không gian, từ đó tích hợp vào hệ thống GIS để xây dựng lớp dữ liệu phục vụ quản lý cây xanh đô thị.

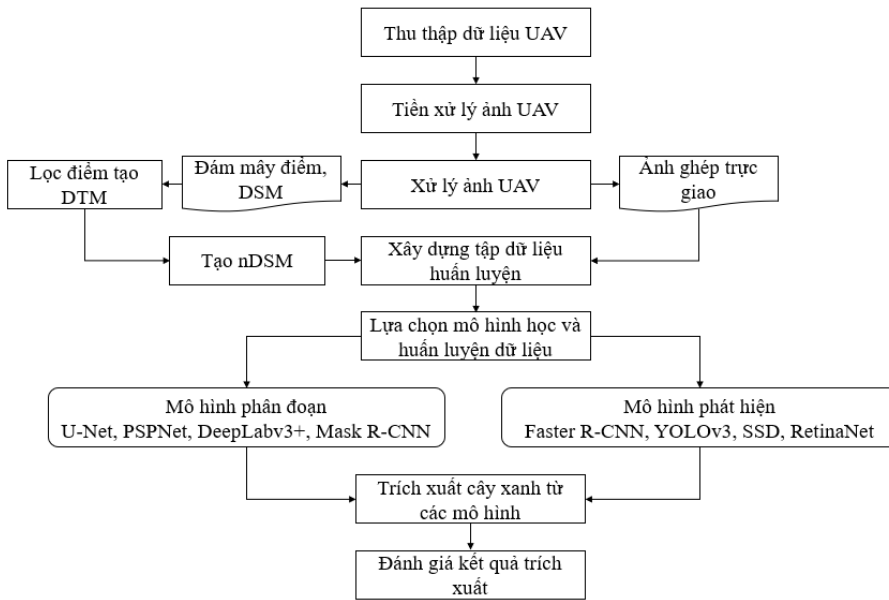
Trên thế giới, học sâu đã được ứng dụng thành công trong nhiều nghiên cứu trích xuất cây xanh. Nổi bật là nghiên cứu của Weinstein và cộng sự (2020) với mô hình DeepForest sử dụng Mask R-CNN để phân đoạn tán cây từ ảnh UAV [9]. Torresan và nhóm nghiên cứu (2017) áp dụng mạng CNN để thành lập bản đồ cây xanh tại khu vực đô thị châu Âu [3]. Gần đây, Neyns và cộng sự (2022) sử dụng DeepLabv3+ để phân đoạn cây xanh từ ảnh vệ tinh, cho thấy hiệu quả cao trong môi trường đô thị đồng đều [10]. Tuy nhiên, phần lớn các nghiên cứu này tập trung vào khu vực rừng hoặc các thành phố có cấu trúc cây xanh tương đối đồng nhất. Trong khi đó, những khu vực như kênh Tẻ tại TP. Hồ Chí Minh, nơi cây xanh xen lẫn với công trình xây dựng và mặt nước, vẫn là thách thức lớn cho mô hình học sâu.

Ngoài ra, việc ứng dụng học sâu tại Việt Nam còn đối mặt với nhiều khó khăn như thiếu bộ dữ liệu huấn luyện chuẩn hóa, điều kiện địa hình và ánh sáng phức tạp, và chưa có quy trình tổng thể tích hợp từ khâu thu thập dữ liệu UAV đến xây dựng lớp thông tin không gian phục vụ công tác quản lý.

Từ thực tiễn đó, nghiên cứu này đề xuất một quy trình ứng dụng kỹ thuật học sâu trong môi trường ArcGIS Pro để trích xuất cây xanh đô thị từ ảnh UAV tại khu vực dọc bờ kênh Tẻ, Quận 7 (cũ). Nghiên cứu tập trung triển khai và so sánh các mô hình học sâu theo hai hướng tiếp cận chính: phân đoạn và phát hiện đối tượng. Các mô hình được đánh giá dựa trên chỉ số Precision nhằm xác định phương án phù hợp nhất cho việc xây dựng lớp dữ liệu không gian hỗ trợ quản lý và quy hoạch cây xanh đô thị.

2. Phương pháp nghiên cứu

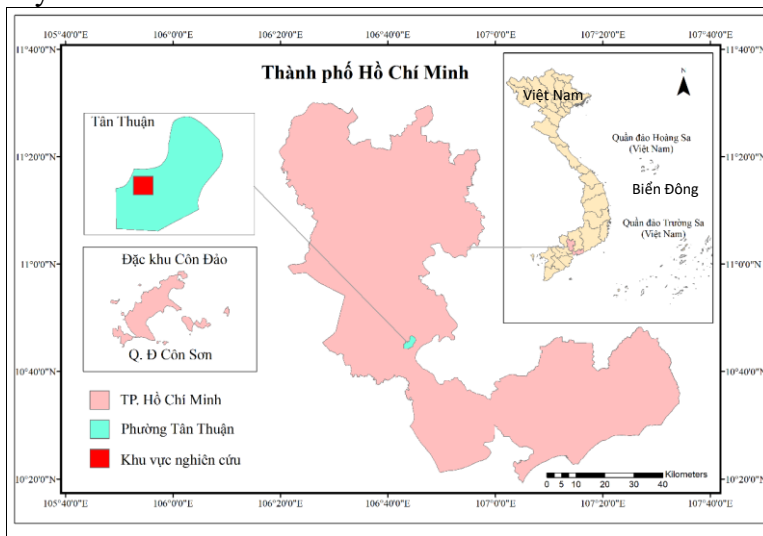
Nghiên cứu này được thực hiện theo một quy trình tổng quát gồm ba giai đoạn chính: (i) thu thập và tiền xử lý ảnh UAV nhằm nâng cao chất lượng hình ảnh đầu vào; (ii) xử lý ảnh để tạo ra các mô hình không gian như DSM, DTM và nDSM; và (iii) ứng dụng mô hình học sâu để nhận dạng cây xanh từ ảnh UAV. Sơ đồ quy trình được trình bày trong Hình 1.



Hình 1. Quy trình thực hiện chung

2.1. Khu vực nghiên cứu

Khu vực nghiên cứu là dọc bờ kênh Tẻ, thuộc phường Tân Thuận, TP. Hồ Chí Minh, diện tích khoảng 14 ha (Hình 2). Kênh Tẻ nối sông Sài Gòn và sông Đòai, với khu dân cư, công viên, và tuyến đường có mật độ cây xanh đa dạng (cây bụi, cây trung bình, cây lớn), xen kẽ công trình và mặt nước. Sự chông lán tán cây và bóng đổ từ tòa nhà tạo thách thức cho trích xuất cây xanh.

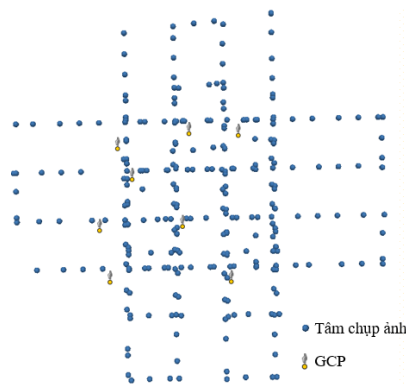


Hình 2. Khu vực nghiên cứu tại phường Tân Thuận, TP. Hồ Chí Minh

2.2. Thu thập và xử lý dữ liệu UAV

Dữ liệu được thu thập bằng UAV DJI Phantom 4 Pro, bao gồm 274 ảnh chụp xiên với 5 góc chụp khác nhau (bốn góc xiên và một góc thẳng đứng) (Hình 3). Dữ liệu được chụp đa góc nhằm phục vụ nhận dạng và trích xuất đối tượng không gian. Độ cao bay chụp là 100 m, với độ phủ dọc và ngang đều là 80%, đảm bảo độ phân giải không gian 2.7 cm/pixel. Tám điểm khống chế mặt đất (Ground Control Point – GCP) được đo bằng kỹ thuật RTK, bố trí đều trong khu vực để nắn chỉnh ảnh, đảm bảo độ chính xác hình học.

Dữ liệu được thu thập trong điều kiện thời tiết ổn định nhằm giảm nhiễu do ánh sáng và đổ bóng thay đổi.



Hình 3. Sơ đồ tuyến bay và vị trí các điểm GCP

Dữ liệu ảnh UAV sau khi thu thập được xử lý sơ bộ nhằm tăng độ tương phản và độ sắc nét, giúp cải thiện hiệu quả phân tích bằng mô hình học sâu. Tiếp theo, quá trình tái tạo mô hình không gian được thực hiện thông qua chuỗi xử lý quang trắc hiện đại. Cụ thể, thuật toán SfM (Structure from Motion) được sử dụng để ước lượng vị trí và hướng của ảnh chụp thông qua việc phát hiện và khớp các điểm đặc trưng, từ đó xây dựng mô hình đám mây điểm thưa (sparse point cloud). Sau đó, kỹ thuật SVM (Multi-View Stereo) được áp dụng để làm dày mật độ điểm, tạo thành đám mây điểm dày đặc (dense point cloud) mô tả chi tiết bề mặt khu vực khảo sát.

Từ đám mây điểm này, mô hình số bề mặt (Digital Surface Model – DSM) được nội suy nhằm phản ánh độ cao tổng thể bao gồm các địa vật như cây xanh, nhà cửa. Để loại bỏ các yếu tố không thuộc mặt đất, đám mây điểm được lọc bằng phần mềm mã nguồn mở Fusion, giữ lại các điểm nằm trên bề mặt địa hình tự nhiên, từ đó nội suy mô hình số địa hình (Digital Terrain Model – DTM). Cuối cùng, mô hình số bề mặt chuẩn hóa (Normalized DSM – nDSM) được tính toán bằng hiệu số giữa DSM và DTM, cung cấp thông tin về chiều cao tuyệt đối của các đối tượng trên mặt đất, đặc biệt hữu ích cho việc nhận dạng cây xanh đô thị.

2.3. Ứng dụng mô hình học sâu để trích xuất cây xanh

2.3.1. Chuẩn bị tập dữ liệu huấn luyện

Tập dữ liệu huấn luyện được xây dựng thông qua quá trình gắn nhãn thủ công các đối tượng cây xanh trên ảnh trực giao RGB, thực hiện trong môi trường ArcGIS Pro. Để hỗ trợ quá trình nhận diện ban đầu, mô hình nDSM được sử dụng để lọc sơ bộ các khu vực tiềm năng có cây xanh, bằng cách chọn các đối tượng có chiều cao lớn hơn 2 m. Bước lọc này giúp loại bỏ các thực thể thấp như cỏ hoặc cây bụi nhỏ, từ đó giảm thiểu thời gian gắn nhãn và tăng độ chính xác trong quá trình chuẩn bị dữ liệu.

Tổng số cây được gắn nhãn ban đầu là 112. Do số lượng mẫu hạn chế có nguy cơ gây ra hiện tượng quá khớp (overfitting) khi huấn luyện mô hình học sâu, kỹ thuật tăng cường dữ liệu được áp dụng để mở rộng tập mẫu và tăng khả năng khái quát hóa của mô hình. Cụ thể, mỗi ảnh được xoay theo các góc 0° , 45° , 90° , 135° , 180° , 225° , 270° và 315° , tạo thành 8 phiên bản khác nhau để tăng tính đa dạng về hình học và góc nhìn – điều đặc biệt quan trọng do đặc điểm tán cây đô thị thường bất đối xứng, bị che khuất hoặc chịu ảnh hưởng bởi bóng đổ.

Mỗi cây được gắn nhãn ở hai dạng: (1) ranh giới khép kín (polygon) dùng cho tác vụ phân đoạn (segmentation), và (2) khung giới hạn chữ nhật (bounding box) dùng cho



tác vụ phát hiện đối tượng (object detection). Dữ liệu sau tăng cường đạt khoảng 800–900 mẫu, đủ để huấn luyện hiệu quả các mô hình học sâu. Tập mẫu ban đầu được chia theo tỷ lệ 70% cho huấn luyện, 20% cho xác thực, và 10% cho kiểm tra. Sau khi huấn luyện, các mô hình sẽ được đánh giá dựa trên kết quả phân đoạn và phát hiện thu được, với thông tin chiều cao cây được trích xuất từ nDSM để bổ sung thuộc tính không gian cho từng đối tượng.

2.3.2. Lựa chọn mô hình học sâu và quy trình huấn luyện

Các mô hình học sâu được sử dụng trong nghiên cứu đều dựa trên kiến trúc CNN, được chia thành hai nhóm chính theo mục tiêu xử lý: phân đoạn (segmentation) và phát hiện (object detection). Nhóm phân đoạn bao gồm các mô hình U-Net, PSPNet, DeepLabv3+ và Mask R-CNN, với chức năng tạo mặt nạ đối tượng nhằm xác định chính xác ranh giới tán cây. Nhóm phát hiện bao gồm Faster R-CNN, YOLOv3, SSD và RetinaNet, được thiết kế để xác định vị trí cây xanh thông qua các khung giới hạn.

Quy trình huấn luyện được thực hiện trong môi trường ArcGIS Pro. Dữ liệu đầu vào được chuẩn hóa và xuất theo định dạng phù hợp, với các tham số huấn luyện được thiết lập nhất quán giữa các mô hình để đảm bảo tính so sánh.

Trong quá trình huấn luyện, hiệu suất của mô hình được đánh giá định kỳ thông qua các chỉ số như giá trị mất mát (loss) và độ chính xác trên tập xác thực (validation accuracy). Các biểu đồ huấn luyện được ghi nhận dưới dạng hình ảnh và nhật ký, hỗ trợ theo dõi diễn tiến huấn luyện và điều chỉnh siêu tham số nhằm tối ưu hóa mô hình. Khi hoàn tất, mô hình có hiệu suất tốt nhất trên tập xác thực sẽ được lưu lại và sử dụng để áp dụng lên tập kiểm tra.

Kết quả đầu ra bao gồm mô hình đã huấn luyện và các lớp đối tượng trích xuất tự động từ ảnh UAV. Các kết quả này tiếp tục được xử lý và đánh giá ở các bước tiếp theo để xác định độ chính xác nhận dạng và mức độ phù hợp cho ứng dụng thực tiễn trong quản lý đô thị.

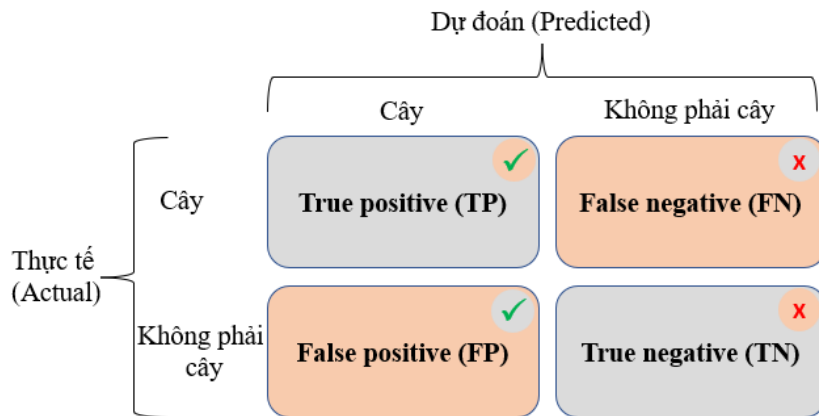
2.4. Đánh giá kết quả và độ chính xác

Đối với nhóm mô hình phân đoạn, sau khi nhận dạng được ranh giới từng tán cây trên ảnh ghép trực giao, chiều cao và độ rộng tán được tính toán nhằm phục vụ phân tích cấu trúc không gian. Chiều cao được trích xuất từ mô hình nDSM tại tâm hình học của vùng phân đoạn, trong khi độ rộng tán được xác định thông qua diện tích mặt nạ hoặc kích thước vùng bao nhỏ nhất. Kết quả này giúp cung cấp thông tin đầy đủ hơn về đặc điểm hình thái của cây xanh, phục vụ quy hoạch và quản lý đô thị.

Ngược lại, các mô hình phát hiện chỉ cung cấp khung giới hạn vị trí của cây nên không đủ cơ sở để tính chiều cao và độ rộng tán. Do đó, nhóm này chủ yếu được sử dụng để thông kê số lượng cây hiện diện trong khu vực nghiên cứu.

Hiệu quả của các mô hình học sâu được đánh giá dựa trên hai tiêu chí chính: độ chính xác (precision) và tổng thời gian thực hiện. Trong đó, độ chính xác phản ánh mức độ tin cậy của mô hình trong việc nhận dạng đúng các đối tượng cây xanh, được tính bằng tỷ lệ giữa số lượng đối tượng được xác định đúng với tổng số đối tượng mà mô hình dự đoán là cây xanh. Chỉ số này đặc biệt quan trọng trong các bài toán phân đoạn ảnh, nhằm đảm bảo rằng các đối tượng được nhận dạng không bị lẫn với các thực thể không liên quan trong ảnh UAV.

Ma trận nhầm lẫn (confusion matrix) như Hình 4 được sử dụng để minh họa rõ mối quan hệ giữa dự đoán của mô hình và thực tế, bao gồm bốn thành phần chính: TP (True Positive), FP (False Positive), FN (False Negative) và TN (True Negative) như minh họa dưới đây:



Hình 4. Ma trận nhầm lẫn

Trong đó, TP là số lượng cây xanh được mô hình nhận dạng đúng; FP là số lượng đối tượng không phải cây nhưng bị nhận nhầm là cây; FN là số cây thật nhưng bị bỏ sót; TN là số đối tượng không phải cây và được mô hình loại bỏ đúng.

Từ ma trận này, công thức tính Precision được xác định như sau:

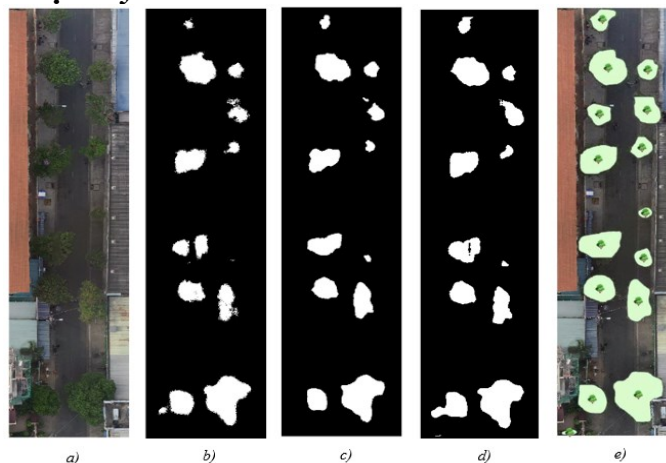
$$Precision = \frac{TP}{TP + FP} \quad (1)$$

Bên cạnh đó, tổng thời gian thực hiện được tính từ thời điểm bắt đầu huấn luyện mô hình cho đến khi hoàn tất quá trình trích xuất kết quả từ ảnh UAV. Tham số này phản ánh mức độ hiệu quả vận hành tổng thể của từng phương pháp, từ đó đánh giá tiềm năng triển khai trong các hệ thống quản lý yêu cầu xử lý thời gian thực.

3. Kết quả nghiên cứu

Các mô hình học sâu được huấn luyện trên tập dữ liệu đã tăng cường và áp dụng trên một ảnh ghép trực giao duy nhất đại diện toàn bộ khu vực nghiên cứu. Kết quả được chia thành hai nhóm chính: phân đoạn và phát hiện, tương ứng với mục tiêu nhận dạng và trích xuất cây xanh phục vụ quản lý đô thị.

3.1. Kết quả phân đoạn cây xanh



Hình 5. Kết quả phân đoạn cây xanh: a) Ảnh gốc; b) U-Net; c) PSPNet; d) DeepLabv3+; e) Mask R-CNN

Bốn mô hình phân đoạn gồm U-Net, PSPNet, DeepLabv3+ và Mask R-CNN đều cho kết quả khả quan trong việc nhận dạng ranh giới tán cây trên ảnh UAV. Trong đó, các mô hình Mask R-CNN và PSPNet đạt độ chính xác cao nhất, với giá trị precision lần lượt là 1.0 và 0.86. Các mô hình DeepLabv3+ và U-Net cũng cho kết quả tốt, với độ chính xác từ 0.79 đến 0.64.

Tổng thời gian thực hiện của các mô hình này dao động từ 15 đến 202 phút, tính từ bước huấn luyện trên tập dữ liệu đã tăng cường đến khi hoàn tất phân đoạn cây trên toàn ảnh ghép. Mask R-CNN yêu cầu thời gian xử lý dài hơn do cấu trúc phức tạp, trong khi U-Net là mô hình đơn giản nhất nên thời gian ngắn hơn.

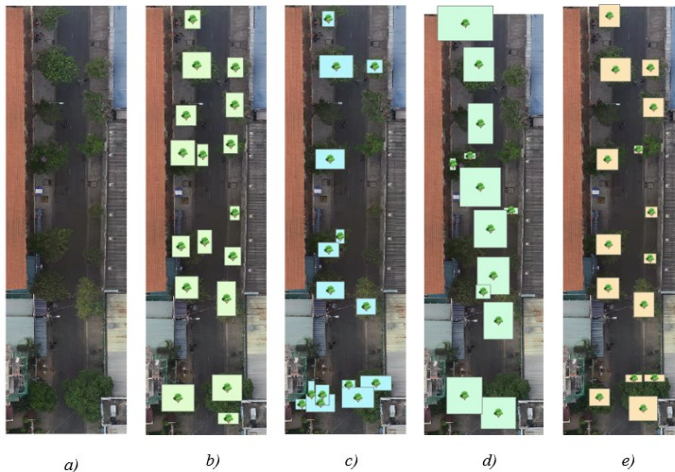
Ngoài kết quả phân loại, các mô hình phân đoạn còn cho phép tính toán chiều cao và độ rộng tán cây dựa trên mặt nạ và dữ liệu nDSM. Điều này góp phần nâng cao giá trị ứng dụng của dữ liệu đầu ra trong các bài toán phân tích không gian và quy hoạch cây xanh. Hình 5 thể hiện kết quả trực quan từ bốn mô hình phân đoạn, với hình a là ảnh gốc và các hình b–e tương ứng với từng mô hình.

3.2. Kết quả phát hiện cây xanh

Với nhóm mô hình phát hiện, Faster R-CNN, YOLOv3, SSD và RetinaNet đều thực hiện được việc xác định vị trí cây xanh bằng khung giới hạn. Trong số đó, RetinaNet đạt độ chính xác cao nhất (precision 0.87), tiếp theo là Faster R-CNN (0.82), SSD (0.69) và YOLOv3 (0.56).

Tổng thời gian thực hiện của các mô hình phát hiện ngắn hơn nhóm phân đoạn, dao động từ 7 đến 47 phút. SSD có thời gian xử lý ngắn nhất, phản ánh đúng đặc trưng của mô hình tối ưu cho tốc độ. Tuy nhiên, các mô hình phát hiện không cung cấp dữ liệu chi tiết về ranh giới hoặc kích thước tán cây, nên không thể tính được các chỉ số hình học bổ sung như chiều cao hay độ rộng tán.

Hình 6 thể hiện kết quả từ bốn mô hình phát hiện, trong đó hình a là ảnh gốc và các hình b–e lần lượt là kết quả từ các mô hình được so sánh.



Hình 6. Kết quả phát hiện cây xanh: a) Ảnh gốc; b) Faster R-CNN; c) YOLOv3; d) SSD; e) RetinaNet

Bảng 1 dưới đây tóm tắt độ chính xác và tổng thời gian thực hiện của từng mô hình, cho thấy nhóm phân đoạn có độ chính xác cao hơn nhưng thời gian xử lý dài hơn, trong khi nhóm phát hiện cho kết quả nhanh hơn nhưng thiếu chi tiết hình học.

Bảng 1. So sánh kết quả giữa các mô hình học sâu

Loại	Mô hình	Precision	Thời gian (phút)
Phân đoạn	U-Net	0.64	15
	PSPNet	0.86	149
	DeepLabv3+	0.79	121
	Mask R-CNN	1.00	202
Phát hiện	Faster R-CNN	0.82	27
	YOLOv3	0.56	47
	SSD	0.69	7
	RetinaNet	0.87	34



4. Thảo luận

Kết quả thực nghiệm cho thấy sự khác biệt đáng kể giữa hai nhóm mô hình học sâu trong trích xuất cây xanh từ ảnh UAV, thể hiện rõ qua độ chính xác và thời gian xử lý. Nhóm mô hình phân đoạn đạt độ chính xác trung bình cao hơn khoảng 12% so với nhóm mô hình phát hiện, xét theo tỷ lệ phần trăm tương đối, đồng thời cung cấp thông tin hình học đầy đủ như ranh giới tán, chiều cao và độ rộng, đây là những yếu tố đặc biệt quan trọng trong quy hoạch và quản lý cây xanh đô thị.

Trong nhóm phân đoạn, Mask R-CNN đạt độ chính xác tuyệt đối (precision = 1.00), cao hơn khoảng 16% so với mô hình xếp sau là PSPNet (precision = 0.86). Sự vượt trội này minh chứng cho khả năng nhận dạng chính xác và đầy đủ ranh giới tán cây của Mask R-CNN. Tuy nhiên, mô hình này cũng có thời gian xử lý dài gấp 13 lần U-Net (202 phút so với 15 phút), phản ánh chi phí tính toán lớn do kiến trúc hai nhánh phức tạp và quy trình xử lý hậu phân đoạn.

Trong nhóm phát hiện, RetinaNet đạt precision = 0.87, cao hơn gần 55% so với YOLOv3 (precision = 0.56). Dù vậy, thời gian xử lý của RetinaNet cũng cao hơn 38%, cho thấy sự đánh đổi giữa độ chính xác và tốc độ vẫn là bài toán then chốt. Đặc biệt, SSD có thời gian xử lý ngắn nhất (7 phút) – nhanh hơn gần 29 lần so với Mask R-CNN, tuy nhiên độ chính xác chỉ đạt 0.69, thấp hơn khoảng 31% so với RetinaNet. Những kết quả này cho thấy các mô hình phát hiện như SSD và YOLOv3 phù hợp hơn trong các tình huống yêu cầu xử lý nhanh hoặc triển khai trên thiết bị hạn chế tài nguyên.

Việc lựa chọn mô hình phù hợp phụ thuộc vào mục tiêu ứng dụng cụ thể. Nếu mục tiêu là xây dựng bản đồ cây xanh có gắn thông tin hình thái phục vụ quản lý chi tiết, nhóm phân đoạn là lựa chọn phù hợp. Trong khi đó, nếu yêu cầu chủ yếu là đếm số lượng hoặc phát hiện nhanh trong môi trường có tài nguyên hạn chế, nhóm phát hiện có thể đáp ứng tốt hơn.

Một điểm hạn chế trong nghiên cứu là số lượng mẫu gắn nhãn ban đầu còn ít dẫn đến khả năng khái quát hóa của mô hình phụ thuộc nhiều vào kỹ thuật tăng cường dữ liệu. Kết quả precision đạt được tuy cao, nhưng vẫn có thể bị ảnh hưởng bởi sự thiếu đa dạng của mẫu huấn luyện. Ngoài ra, nghiên cứu mới chỉ đánh giá mô hình trên một ảnh ghép duy nhất. Để tăng tính tin cậy, các nghiên cứu tiếp theo cần áp dụng trên tập ảnh đa khu vực và sử dụng kỹ thuật đánh giá chéo.

Cuối cùng, mặc dù ArcGIS Pro hỗ trợ tích hợp trọn bộ quy trình huấn luyện và áp dụng mô hình học sâu một cách thuận tiện, nền tảng này vẫn còn những hạn chế nhất định về khả năng tùy biến thuật toán, cấu hình kiến trúc mạng và kiểm soát chi tiết quá trình huấn luyện. Trong nghiên cứu này, việc huấn luyện mô hình được thực hiện hoàn toàn trong môi trường ArcGIS Pro, do đó chưa có cơ sở để so sánh trực tiếp với các nền tảng mã nguồn mở như PyTorch hoặc TensorFlow. Tuy nhiên, các nền tảng lập trình này vốn được biết đến với tính linh hoạt cao và cộng đồng phát triển mạnh, cho phép tinh chỉnh sâu hơn về siêu tham số và cấu trúc mạng. Do vậy, việc mở rộng nghiên cứu sang các nền tảng này trong tương lai có thể là một hướng đi triển vọng nhằm kiểm định lại hiệu suất mô hình, đồng thời khai thác các kiến trúc mới tiên tiến hơn để cải thiện chất lượng trích xuất thông tin không gian từ ảnh UAV.

5. Kết luận

Nghiên cứu đã đề xuất và triển khai một quy trình trích xuất cây xanh đô thị từ ảnh UAV bằng kỹ thuật học sâu tích hợp trong ArcGIS Pro, với trọng tâm là so sánh hiệu quả giữa các mô hình phân đoạn và phát hiện đối tượng. Phương pháp đã tận dụng dữ liệu ảnh



ghép trực giao và mô hình nDSM để hỗ trợ gắn nhãn và tiền xử lý, đảm bảo tính khả thi trong môi trường đô thị phức tạp.

Kết quả thực nghiệm xác nhận rằng nhóm mô hình phân đoạn có ưu thế trong việc cung cấp dữ liệu chi tiết về ranh giới và hình thái tán cây, trong khi nhóm mô hình phát hiện phù hợp hơn cho các ứng dụng yêu cầu thời gian xử lý nhanh. Nghiên cứu góp phần khẳng định vai trò của học sâu trong xây dựng dữ liệu không gian cây xanh phục vụ quản lý đô thị, đặc biệt khi được kết hợp với nền tảng UAV và hệ thống GIS.

Hướng phát triển tiếp theo sẽ tập trung vào mở rộng quy mô vùng nghiên cứu, đa dạng hóa tập dữ liệu huấn luyện và tích hợp các mô hình tiên tiến để nâng cao tính tổng quát và khả năng ứng dụng vào thực tiễn.

Lời cảm ơn

Các tác giả xin trân trọng cảm ơn Trường Đại học Bách Khoa TP. Hồ Chí Minh, Trường Đại học Tài nguyên và Môi trường TP. Hồ Chí Minh đã tạo điều kiện thuận lợi về cơ sở vật chất và môi trường nghiên cứu trong suốt quá trình thực hiện và hoàn thiện bài báo này.

Cam kết của các tác giả

Tất cả các tác giả có tên trong bài báo cam kết đã thống nhất và đồng thuận trong việc công bố các kết quả nghiên cứu, đồng thời không có bất kỳ xung đột lợi ích nào liên quan đến nội dung được trình bày trong bài báo này.

Tài liệu tham khảo

- [1] Nowak, David J., Hirabayashi, Satoshi, Bodine, Allison, Greenfield, Eric, *Tree and forest effects on air quality and human health in the United States*. Environmental Pollution, 2014. 193: p. 119-129.
- [2] National Parks Board-Singapore, *Managing Our City in Nature: Technology-Driven Urban Greenery Solutions*. 2023.
- [3] Torresan, Chiara. , Berton, Andrea. , Carotenuto, Federico. , Gennaro, Salvatore Filippo Di. , Gioli, Beniamino. , Matese, Alessandro. , Miglietta, F. , Vagnoli, Carolina. , Zaldei, Alessandro. , Wallace, Luke Oliver. , *Forestry applications of UAVs in Europe: A review*. International Journal of Remote Sensing 2016. 38(8-10): p. 1-21.
- [4] VnExpress. Hơn 200.000 cây ở TP HCM được kiểm tra, tránh gãy đổ ra sao? Truy cập ngày 9/6/2025; Available from: <https://vnexpress.net/hon-200-000-cay-o-tp-hcm-duoc-kiem-tra-tranh-gay-do-ra-sao-4780326.html>.
- [5] Trần Thị Vân, Phạm Khánh Hòa, Thảm Thị Ngọc Hân, *Đánh giá thực trạng không gian xanh - Thuộc đo chất lượng môi trường hướng đến phát triển đô thị xanh cho Thành phố Hồ Chí Minh*. Journal of Transportation Science and Technology, 2018. 29: p. 56-63.
- [6] Trần Ngọc Huyền Trang, Trần Thị Kim Oanh, Lê Trung Chon, *Kết hợp thiết bị bay không người lái (UAV) và GIS xây dựng dữ liệu phục vụ quản lý cây xanh đô thị*. Vietnam Journal of Hydrometeorology, 2022. EME4(1): p. 139-148.
- [7] Nguyễn Quốc Tuấn, *Ứng dụng GIS mã nguồn mở trong quản lý cây xanh đô thị tại phường 6, quận 3, Thành phố Hồ Chí Minh*. 2011, Trường Đại học Nông Lâm. p. 81.
- [8] Yann, LeCunY., BengioY., Geoffrey, Hinton., *Deep Learning*. Nature, 2015. 521(7553): p. 436-444.
- [9] Weinstein, Ben. , Marconi, Sergio. , Méline, Aubry-Kientz., Vincent, Gregoire. , Senyondo, Henry. , White, Ethan P. , *DeepForest: A Python package for RGB image-based individual tree crown detection*. Methods in Ecology and Evolution, 2020. 11(12): p. 1741-1752.
- [10] Robbe, Neyns., Frank, Canters., *Mapping of Urban Vegetation with High-Resolution Remote Sensing: A Review*. Remote Sensing, 2022. 14(4).

Article © 2024 by Magazine of Geodesy - Cartography is licensed under CC BY 4.0

