

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC MỎ - ĐỊA CHẤT

BÁO CÁO HỌC THUẬT

NĂM HỌC 2024-2025

Học kỳ 2

**Quy trình trích xuất đối tượng từ ảnh UAV sử dụng
mô hình học sâu**

Người báo cáo: TS Phạm Trung Dũng

Đơn vị: Bộ môn Trắc địa Công trình

Hà Nội, Tháng 6 năm 2025

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC MỎ - ĐỊA CHẤT

BÁO CÁO HỌC THUẬT

NĂM HỌC 2024-2025

Học kỳ 2

**Quy trình trích xuất đối tượng từ ảnh UAV sử dụng
mô hình học sâu**

Phòng KHCN

Bộ môn

Người báo cáo

TS Phạm Trung Dũng

Hà Nội, tháng 6 năm 2025

MỤC LỤC

MỞ ĐẦU	3
Chương 1 XÂY DỰNG QUY TRÌNH TRÍCH XUẤT DỮ LIỆU TỪ ẢNH UAV	4
<i>1.1 Thu thập dữ liệu từ thiết bị bay không người lái UAV</i>	4
<i>1.2 Tạo mẫu phục vụ quá trình đào tạo sử dụng mạng học sâu</i>	12
<i>1.3 Quy trình trích xuất dữ liệu ảnh sử dụng mạng học sâu</i>	18
Chương 2 THỰC NGHIỆM	21
2.1 Dữ liệu khu vực nghiên cứu	21
<i>2.1.1 Tập dữ liệu huấn luyện mô hình (Training data set)</i>	21
<i>2.1.2 Tập dữ liệu đánh giá hiệu năng mô hình học máy trong quá trình huấn luyện (Validation data set)</i>	21
<i>2.1.3 Tập dữ liệu đánh giá hiệu năng mô hình học máy sau quá trình huấn luyện (Testing data set)</i>	22
2.3 Phát triển mô hình Unet để phân loại đối tượng nhà	23
2.4 Kết quả dự đoán (trích xuất) lớp nhà từ mô hình	24
KẾT LUẬN VÀ KIẾN NGHỊ	30
TÀI LIỆU THAM KHẢO	31

DANH MỤC HÌNH VẼ, ĐỒ THỊ

Hình 1-1: Sơ đồ quy trình xử lý ảnh UAV	5
Hình 1-2: Xác định ranh giới khu vực bay chụp	6
Hình 1-3: Ví dụ về đo điểm khống chế tọa độ và độ cao bằng công nghệ GNSS tại khu vực Bình Sơn, tỉnh Quảng Ngãi.....	7
Hình 1-4: Trạm điều khiển mặt đất GNSS (trái); Tiến hành bay UAV (phải)	8
Hình 1-5: Giao diện Coprocess.....	9
Hình 1-6: Point cloud khu vực dự án.....	10
Hình 1-7: Công cụ kiểm tra độ cao bằng phần mềm	10
Hình 1-8: Số liệu kiểm tra độ cao của phần mềm.....	10
Hình 1-9: Đưa ảnh trực giao trên phần mềm Global Mapper	11
Hình 1-10: Đo tọa độ của các điểm kiểm tra	11
Hình 1-11: Ảnh trực giao sau khi nắn trên phần mềm QGIS	12
Hình 1-12: Tạo lớp dữ liệu số hóa	12
Hình 1-13: Chọn kiểu topology cho đối tượng cần số hóa	13
Hình 1-14: Chọn hệ quy chiếu.....	13
Hình 1-15: Polygon của đối tượng nhà sau khi số hóa trên ảnh UAV	14
Hình 1-16: Kiểm tra topology của đối tượng đã số hóa	14
Hình 1-17: Xuất file dữ liệu đã số hóa sang dạng shapefile sử dụng phần mềm QGIS	15
Hình 1-18: Các mẫu tiles được chia nhỏ từ một tấm ảnh UAV ortho lớn	15
Hình 1-19: Các mask tương ứng với các tiles.....	16
Hình 1-20: Các đối tượng là Nhà được vector hóa tương ứng với các tiles	17
Hình 1-21: Một mẫu từ tập dữ liệu trong khu vực nghiên cứu trong đó (a)-ảnh trực giao, (b)- là vector của tòa nhà, và (c) là mask được raster hóa tòa nhà.	18
Hình 1-22: Trích xuất đối tượng nhà từ ảnh RGB+ nDSM trên mô hình học sâu với đầu ra là các polygon (Sun 2021)	19
Hình 1-23: Sơ đồ quy trình trích xuất dữ liệu sử dụng mạng học sâu	20
Hình 0-1: Ảnh UAV sau khi nắn khu vực kiểm tra testing	27
Hình 0-2: Lớp nhà được số hóa (vector) sử dụng làm đối tượng thực để testing	27
Hình 0-3: Kết quả dự đoán từ mô hình Unet	28
Hình 0-4: So sánh kết quả dự đoán (đường màu đỏ) và đối tượng thực (màu vàng)	28

MỞ ĐẦU

Tự động trích xuất đối tượng trực tiếp từ ảnh chụp của máy bay không người lái (UAV) là vấn đề đang thu hút được sự quan tâm lớn không chỉ của các nhà nghiên cứu mà cả các cơ sở sản xuất trong lĩnh vực trắc địa bản đồ. Tự động trích xuất đối tượng trên ảnh UAV là quá trình nhận dạng và véc tơ hóa đối tượng trực tiếp từ ảnh chụp UAV. Các địa vật như lớp giao thông, thủy hệ, địa vật và trong đó có đối tượng nhà là một trong những đối tượng quan trọng cần được thể hiện trong công tác thành lập và cập nhật các loại bản đồ địa hình, địa chính và bản đồ quy hoạch phục vụ công tác giám sát xây dựng, quản lý và lập kế hoạch các công trình nói chung và công trình cảng biển nói riêng. Do đó, trích xuất đối tượng trên ảnh UAV đang là nhu cầu cần thiết trong công tác biên tập bản đồ. Hiện nay, nhờ sự phát triển mạnh mẽ của thị giác máy tính (computer vision) và công nghệ mạng nơ-ron nhận tạo học sâu (deep learning) đã mở ra cánh cửa cho khả năng hiện thực hóa việc trích xuất đối tượng trực tiếp và tự động trên ảnh UAV.

Những năm gần đây ở Việt nam, công nghệ UAV đã dần trở thành một trong những công cụ phổ biến nhất trong lĩnh vực đo đạc bản đồ. Đặc biệt kể từ sau thông tư số 7/2021 của Bộ Tài nguyên và Môi trường ban hành quy định kỹ thuật sử dụng UAV để cập nhật cơ sở dữ liệu nền địa lý quốc gia và thành lập bản đồ địa hình tỉ lệ lớn . Bởi vì, công nghệ UAV có nhiều ưu điểm vượt trội so với công nghệ truyền thống như tốc độ hoàn thành công việc, giảm chi phí giá thành sản phẩm, dễ dàng và nhanh chóng triển khai trong những khu vực khó tiếp cận (ví dụ: khu vực núi cao, khu vực đang tranh chấp...). Cho đến nay, đa số các bước chính trong quy trình thành lập bản đồ từ ảnh UAV như: thiết kế bay, nắn ảnh, tạo mô hình số độ cao, tạo ảnh trực giao đều có thể tự động nhờ các phần mềm chuyên dụng. Tuy nhiên, hiện nay chưa có quy trình quy phạm và các tài liệu hướng dẫn công tác trích xuất dữ liệu trên tư liệu ảnh sử dụng mạng học sâu. Nhận thấy được vấn đề tồn tại nêu trên nội dung nghiên cứu được đề xuất: **“Quy trình trích xuất đối tượng từ ảnh UAV sử dụng mô hình học sâu”**.

Chương 1 XÂY DỰNG QUY TRÌNH TRÍCH XUẤT DỮ LIỆU TỪ ẢNH UAV

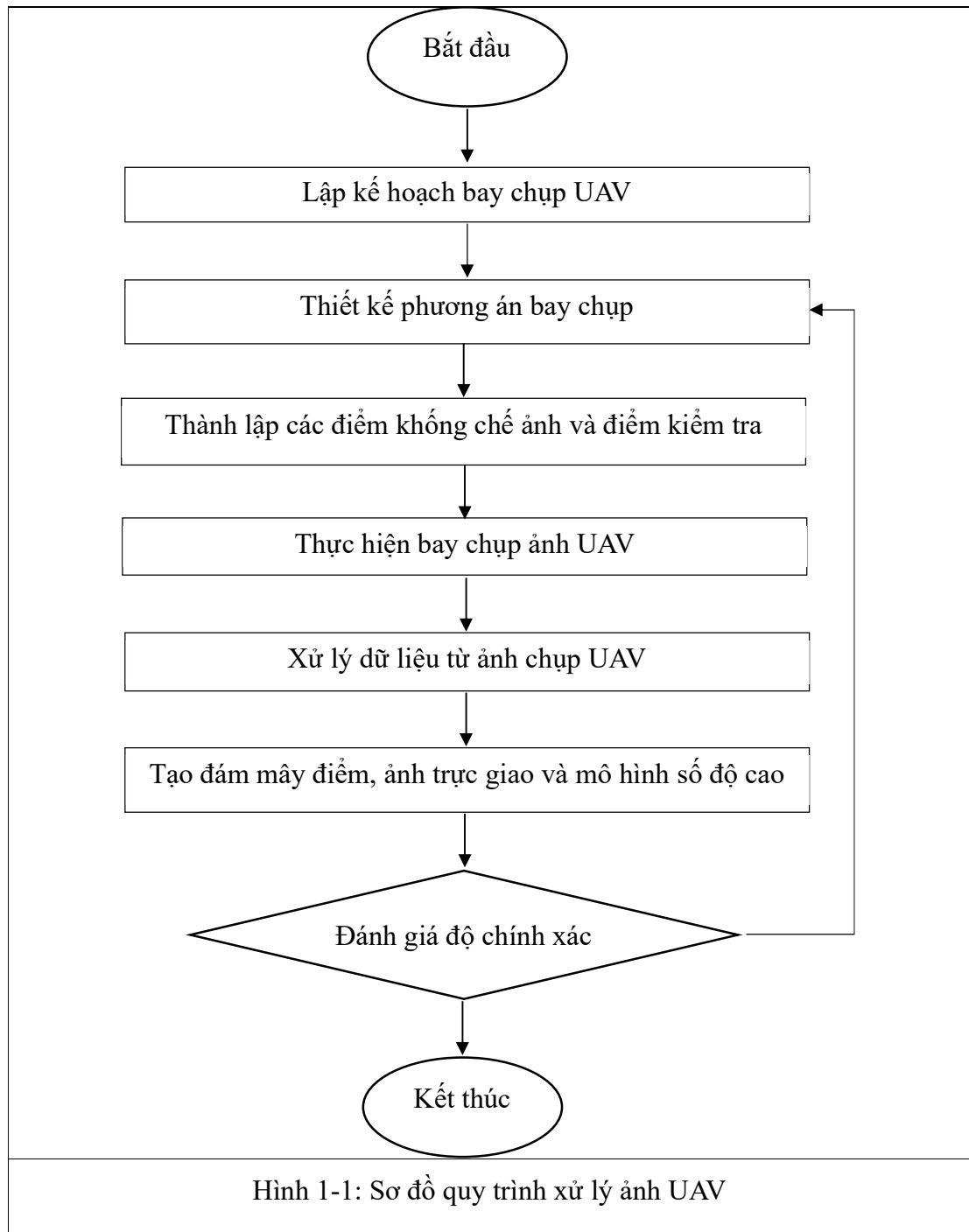
Trích xuất dữ liệu từ ảnh UAV để được các đối tượng cần thiết và thể hiện trên bản đồ địa hình, bản đồ địa chính cần phải thực hiện theo một quy trình chặt chẽ và phù hợp. Mặc dù ở nước ta đã ban hành thông tư số 07/2021 về việc xây dựng bản đồ địa hình tỉ lệ lớn từ dữ liệu UAV tuy nhiên các công tác trích xuất dữ liệu từ ảnh UAV phần lớn vẫn phải thực hiện thủ công. Kế thừa những thành tựu phát triển của cuộc cách mạng trí tuệ nhân tạo mà hiện nay việc trích xuất thông tin các đối tượng trên ảnh đã khả thi và đang là xu hướng. Những lợi ích to lớn từ việc trích xuất dữ liệu tự động như giảm giá thành, tăng năng suất lao động và đáp ứng kịp thời cho sản xuất là lý do để nhiều nước đầu tư nghiên cứu. Tuy nhiên ở nước ta, công tác trích xuất dữ liệu tự động từ ảnh UAV để thành lập bản đồ và cập nhật dữ liệu vẫn còn là một chủ đề mới. Do đó, việc nghiên cứu để ban hành các hướng dẫn kỹ thuật, quy trình trích xuất dữ liệu tự động từ thiết bị UAV một cách chặt chẽ và phù hợp với điều kiện sản xuất ở nước ta là một hướng đi cần thiết.

Trong chương này sẽ giới thiệu cấu trúc dữ liệu ảnh UAV và tiến hành đề xuất quy trình tự động trích xuất dữ liệu sử dụng kỹ thuật của mạng học sâu từ ảnh UAV.

Quy trình được đề xuất bao gồm bốn bước cơ bản: (1) Thu thập dữ liệu mặt đất từ ảnh UAV; (2) Tạo mẫu dữ liệu cho mô hình học máy; (3) Xây dựng mô hình; và (4) trích xuất đối tượng tự động và trực tiếp từ ảnh UAV.

1.1 Thu thập dữ liệu từ thiết bị bay không người lái UAV

Việc thu thập dữ liệu từ ảnh UAV phục vụ cho công tác thành lập bản đồ đã được thực hiện ở nước ta từ nhiều năm qua. Có nhiều nghiên cứu và các hướng dẫn kỹ thuật về quy trình tạo ra ảnh trực giao (ortho images), mô hình số độ cao (DEM), và dữ liệu đám mây điểm (point cloud) từ ảnh chụp UAV (Lê, Lê et al. 2020), (Long, Goyal et al. 2021), (Long and Cường 2019), (Sỹ, Quý et al. 2017), (Cường, Trường et al. 2019). Tuy nhiên, để thành lập ảnh trực giao, mô hình số độ cao và đám mây điểm từ ảnh chụp UAV có thể tóm tắt lại thông qua các bước chính như sơ đồ Hình 1-1.



1. Lập kế hoạch bay chụp UAV

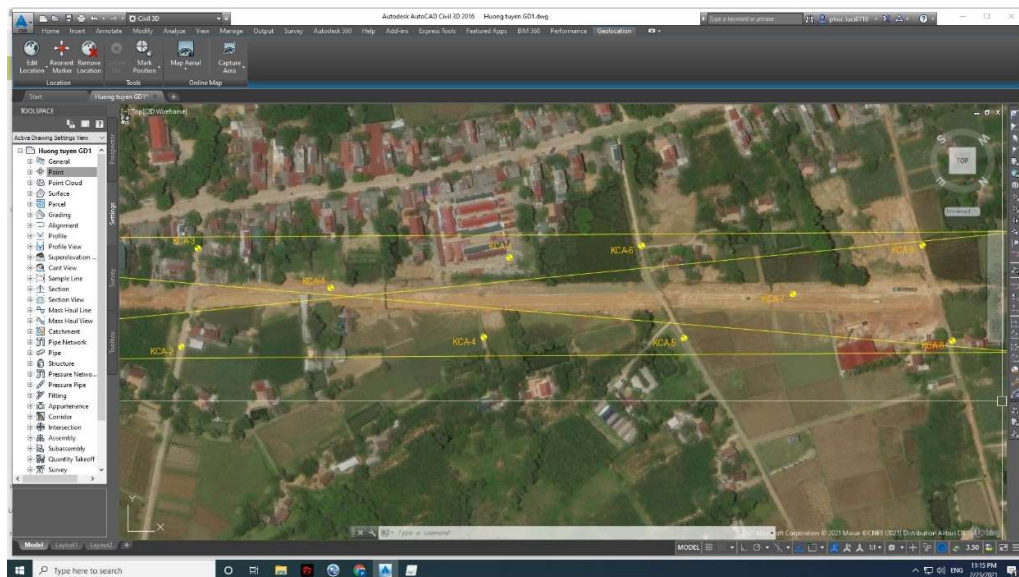
Dựa trên thông tư 07/2021/TT- BTNMT về “Quy định kỹ thuật thu nhận và xử lý dữ liệu ảnh số từ tàu bay không người lái phục vụ xây dựng, cập nhật cơ sở dữ liệu nền địa lý quốc gia tỷ lệ 1:2.000, 1:5.000 và thành lập bản đồ địa hình tỷ lệ 1:500, 1:1.000. Ngoài ra, xem xét các quy định, quy phạm trong thành lập bản đồ địa hình với các

đặc điểm địa lý, kinh tế - xã hội của khu vực thành lập bản đồ. Sau đó tiến hành các công việc sau để lên kế hoạch bay chụp UAV:

- Liên hệ với các cơ quan nhà nước, chính quyền địa phương làm thủ tục và xin giấy phép bay
- Chuẩn bị số liệu tọa độ, độ cao các điểm khống chế các cấp
- Khảo sát khu đo, chọn vị trí cất hạ cánh: Vị trí cất hạ cánh là các bãi trống trong khu vực đảm bảo cách xa khu dân cư, các công trình nhà cao tầng, các cột phát sóng...
- Chuẩn bị và kiểm tra các thiết bị máy và phương tiện phục vụ công tác bay chụp ảnh.

2. Thiết kế phương án bay chụp

Dựa trên tọa độ các điểm ranh giới của khu vực khảo sát tiến hành tính chuyển sang tọa độ WGS-84 và xuất ra file KML, KMZ và trực quan hóa khu đo trên ứng dụng phần mềm Google Earth. Tiếp theo, tiến hành xác định chiều cao bay và mức độ độ chồng phủ (dọc và ngang) của ảnh, từ đó tính toán số ca bay với lưu ý đến thời lượng pin/1 ca (trung bình 15phút/1 ca bay đối với Phantom 4 RTK).



Hình 1-2: Xác định ranh giới khu vực bay chụp

3. Thành lập các điểm khống chế ảnh và điểm kiểm tra

Tiến hành tìm kiếm các điểm mốc khống chế trắc địa mặt bằng và độ cao đã tồn tại trong khu vực và đánh giá khả năng sử dụng của các điểm này. Tiếp theo, tiến hành chèn dày hoặc thành lập mới các điểm khống chế và điểm kiểm tra tuân thủ theo các quy phạm hiện hành. Các điểm khống chế và điểm kiểm tra phục vụ công tác bay

chụp cần đảm bảo cả về độ chính xác cũng như sự thông thoáng và rõ nét cho mục đích bay chụp UAV.



Hình 1-3: Ví dụ về đo điểm khống chế tọa độ và độ cao bằng công nghệ GNSS tại khu vực Bình Sơn, tỉnh Quảng Ngãi

4. Thực hiện bay chụp ảnh UAV

Sau khi thiết kế ca bay và thành lập các điểm khống chế cũng như các điểm kiểm tra tiến hành bay chụp UAV. Công tác bay chụp UAV cần phải đảm bảo các yêu cầu sau:

- Kiểm tra đầy đủ pin cho thiết bị UAV, bộ phận điều khiển từ xa RC và bộ máy phát tín hiệu GNSS từ trạm base mặt đất.
- Cài đặt trạm Base theo phương pháp PPK để thu thập dữ liệu đo tĩnh tại mốc khống chế tọa độ và độ cao trong khu vực bay chụp.
- Kiểm tra thẻ SD và dung lượng bộ nhớ SD trong thiết bị bay.
- Tiếp theo tiến hành công tác bay:
 - Khởi động trạm điều khiển và phần mềm điều khiển thiết bị bay
 - Kiểm tra kết nối giữa tay điều khiển và màn hình kiểm soát
 - Kiểm tra tín hiệu GNSS và tín hiệu kết nối với trạm điều khiển
 - Thực hiện cất cánh thiết bị bay
 - Theo dõi quá trình hoạt động của thiết bị bay như: tín hiệu kết nối với điều khiển, mức pin, tốc độ gió và các cảnh báo khác đến khi kết thúc ca bay
 - Kiểm tra dữ liệu thu nhận và lưu dữ liệu từ thẻ nhớ sang máy tính hoặc các thiết bị khác.



Hình 1-4: Trạm điều khiển mặt đất GNSS (trái); Tiến hành bay UAV (phải)

5. Xử lý dữ liệu từ ảnh chụp UAV

Ảnh sau khi thu nhận được bằng thiết bị bay không người lái sẽ phải thực hiện nắn ảnh thông qua hai bài toán cơ bản là định hướng trong (định hướng tương đối) và định hướng ngoài (định hướng tuyệt đối). Các ảnh được định hướng tương đối nhờ độ chồng phủ dọc và ngang giữa các tấm ảnh khi chụp. Để đưa tọa độ mô hình về hệ tọa độ trắc địa cần sử dụng các điểm khống chế ảnh mặt đất hoặc sử dụng tọa độ tâm chụp ảnh. Do đó để xử lý dữ liệu từ ảnh chụp UAV cần phải thực hiện hai công việc chính là xác định tọa độ tâm chụp và xử lý ảnh được trình bày chi tiết như sau:

a) Xác định tọa độ tâm chụp ảnh

Đối với trường hợp sử dụng thiết bị bay có bộ định vị RTK thì thực hiện xác định tâm chụp ảnh, trường hợp này sẽ giảm đáng kể số lượng điểm khống chế ảnh mặt đất. Trường hợp sử dụng thiết bị bay Phantom 4 RTK và phần mềm CHC CGO 2.0 thì các bước thực hiện theo các bước sau:

b) Xử lý ảnh UAV

Ảnh chụp từ thiết bị bay không người lái được xử lý bằng phần mềm Agisoft Metashape Professional.

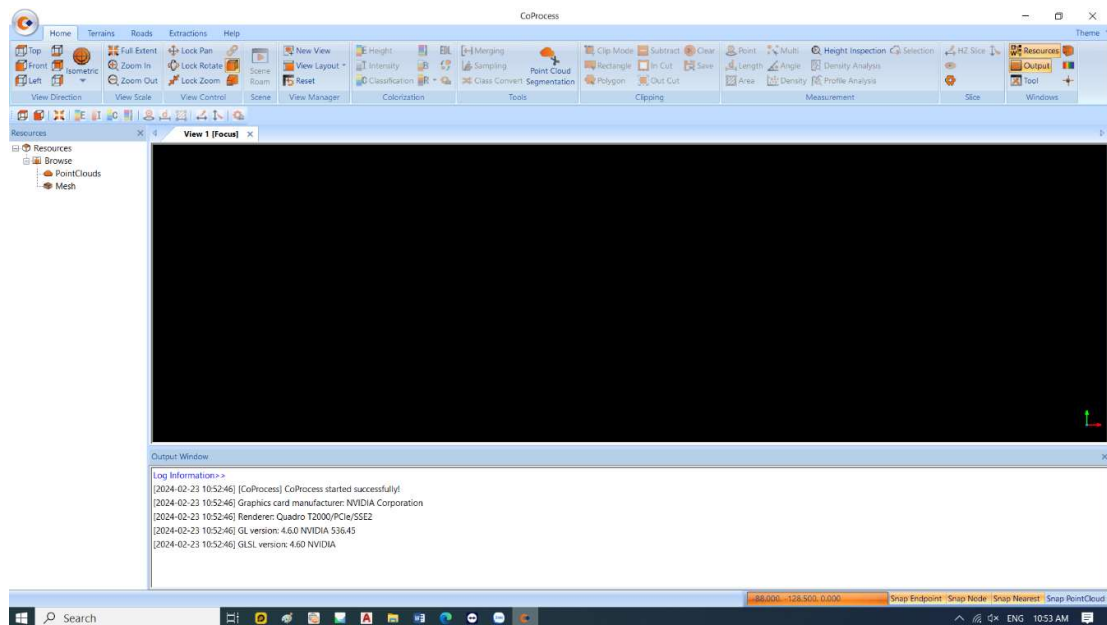
6 Đánh giá độ chính xác

Quá trình xử lý ảnh cần phải được kiểm tra đánh giá để bảo đảm sản phẩm giao nộp đạt các yêu cầu như đã đề ra thông qua công tác đánh giá độ chính xác. Đánh giá độ chính xác được thực hiện dựa trên nguyên tắc so sánh tọa độ và độ cao của các điểm được đo trên ảnh và thực địa. Sau đây sẽ giới thiệu các bước đánh giá độ chính xác độ cao sử dụng phần mềm CHCNAV và đánh giá độ chính xác về mặt bằng sử dụng phần mềm Global Mapper.

a. Đánh giá độ chính xác về độ cao

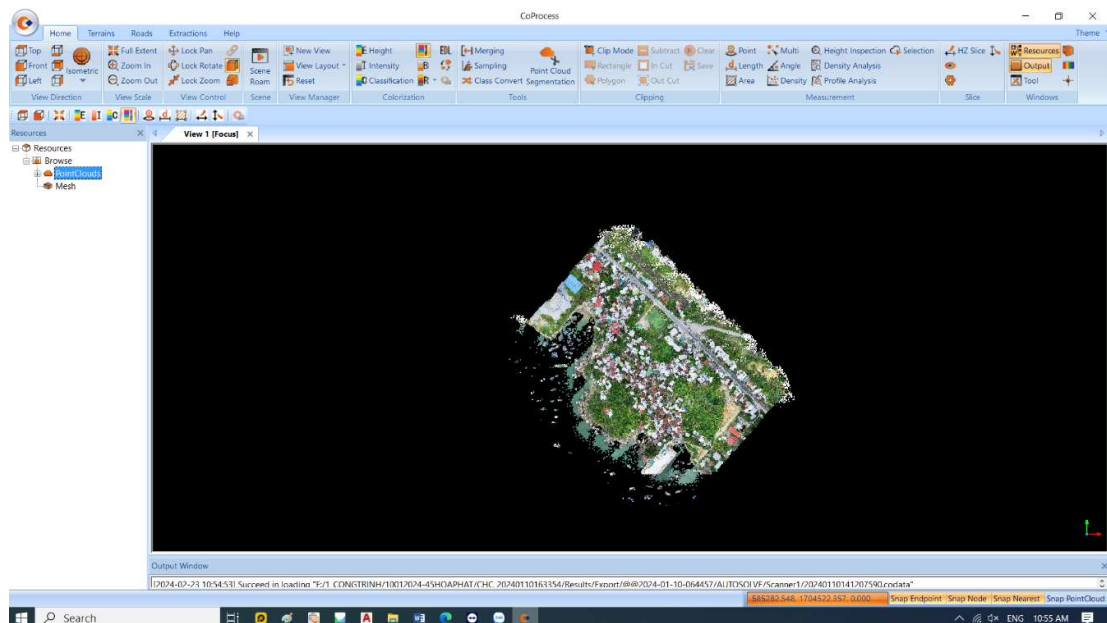
Đánh giá độ chính xác độ cao bằng phần mềm CHCNAV dựa trên nguyên tắc so sánh độ cao của điểm kiểm tra đo trên thực địa và các điểm lân cận gần nhất từ đám mây điểm được xác định trên ảnh. Sau đây là các bước thực hiện bởi phần mềm:

Bước 1: Khởi động phần mềm Coprocess (CHCNAV)



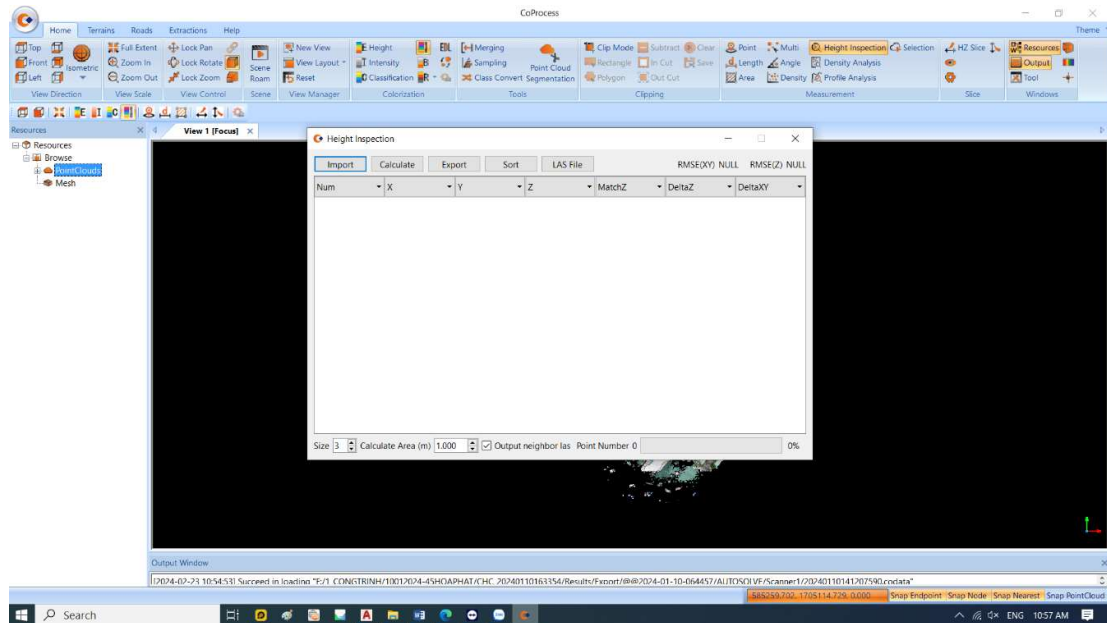
Hình 1-5: Giao diện Coprocess

Bước 2: Đưa số liệu Point Cloud vào phần mềm



Hình 1-6: Point cloud khu vực dự án

Bước 3: Đưa số liệu đo bằng máy GNSS RTK i83 So sánh tọa độ, cao độ các điểm kiểm tra.



Hình 1-7: Công cụ kiểm tra độ cao bằng phần mềm

Height Inspection							
				RMSE(XY) 0.130		RMSE(Z) 0.070	
Num	X	Y	Z	MatchZ	DeltaZ	DeltaXY	
A8	586086.656	1704524.532	20.390	20.472	-0.082	0.590	
A9	586085.194	1704513.835	20.550	20.629	-0.079	0.035	
A10	586049.484	1704473.730	23.220	23.238	-0.018	0.103	
A11	586032.387	1704473.841	24.670	24.692	-0.022	0.062	
A12	586022.688	1704481.535	25.390	25.456	-0.066	0.098	
A13	585961.009	1704501.310	26.430	26.484	-0.054	0.096	
A14	585945.691	1704482.245	25.880	25.967	-0.087	0.105	
A15	585944.799	1704404.470	29.240	29.277	-0.037	0.067	
A16	585949.716	1704430.236	27.570	27.643	-0.073	0.051	
A17	585938.009	1704459.815	20.670	20.689	0.019	0.033	
A18	585974.653	1704459.272	26.430	26.496	-0.066	0.044	
A19	586211.527	1704581.097	22.350	22.430	-0.080	0.055	

Hình 1-8: Số liệu kiểm tra độ cao của phần mềm

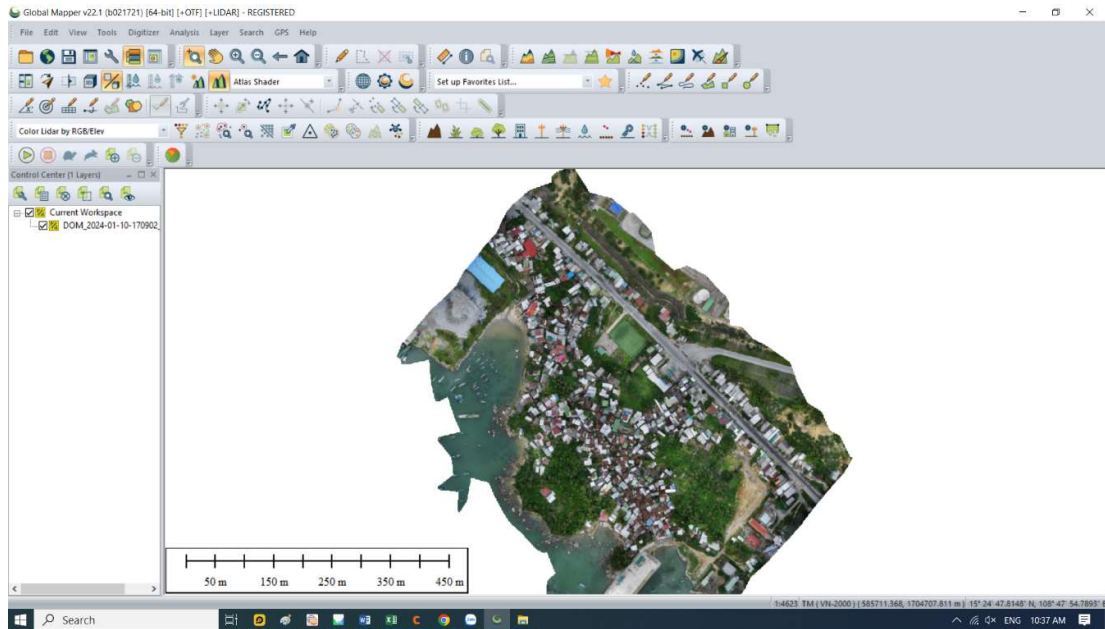
Kết quả sai số độ cao với 19 điểm đạt 7 cm cho thấy công nghệ UAV hoàn toàn đáp ứng được yêu cầu độ chính xác độ cao trong công tác thành lập bản đồ tỉ lệ lớn.

b. Đánh giá độ chính xác về tọa độ

Đánh giá độ chính xác mặt bằng sử dụng phần mềm Global Mapper dựa trên nguyên tắc so sánh tọa độ của các điểm kiểm tra có hình rõ nét trên ảnh và đã được xác định

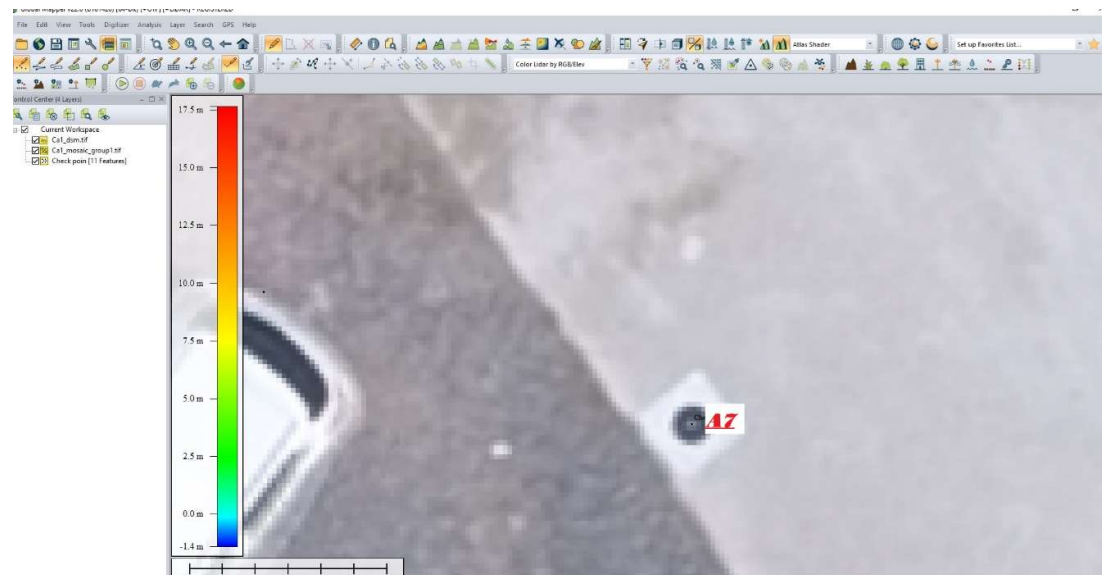
tọa độ ở thực địa. Việc đánh giá tiến hành bằng cách so sánh tọa độ trên ảnh và tọa độ thực tế đã xác định bằng GNSS hoặc toàn đạc điện tử ở ngoài thực địa. Sau đây sẽ giới thiệu các bước cụ thể trên phần mềm Global Mapper:

Bước 1: Đưa ảnh trực giao Orthomosaic vào phần mềm Global Mapper



Hình 1-9: Đưa ảnh trực giao trên phần mềm Global Mapper

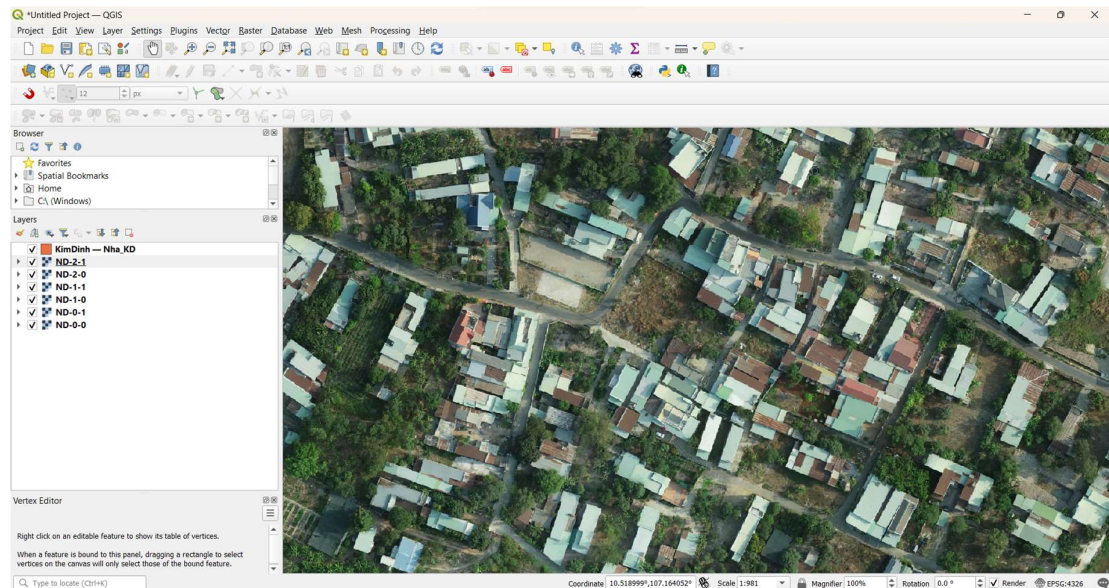
Bước 2: Đo tọa độ của các điểm kiểm tra trên ảnh



Hình 1-10: Đo tọa độ của các điểm kiểm tra

1.2 Tạo mẫu phục vụ quá trình đào tạo sử dụng mạng học sâu

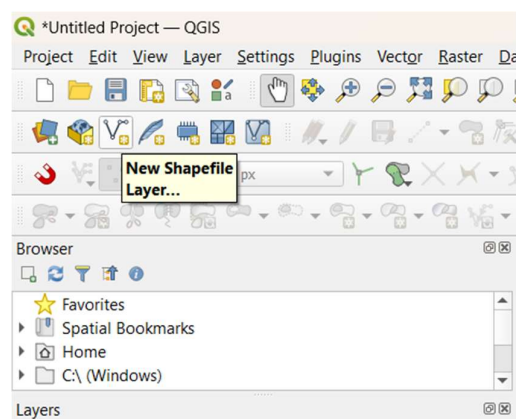
Từ ảnh sau khi đã nắn (ảnh trực giao - ortho) sẽ được sử dụng để vector hóa (số hóa) các đối tượng nhà trực tiếp trên ảnh. Có nhiều công cụ cho phép sử dụng để số hóa lớp đối tượng nhà như AutoCad, Arcgis, QGIS...



Hình 1-11: Ảnh trực giao sau khi nắn trên phần mềm QGIS

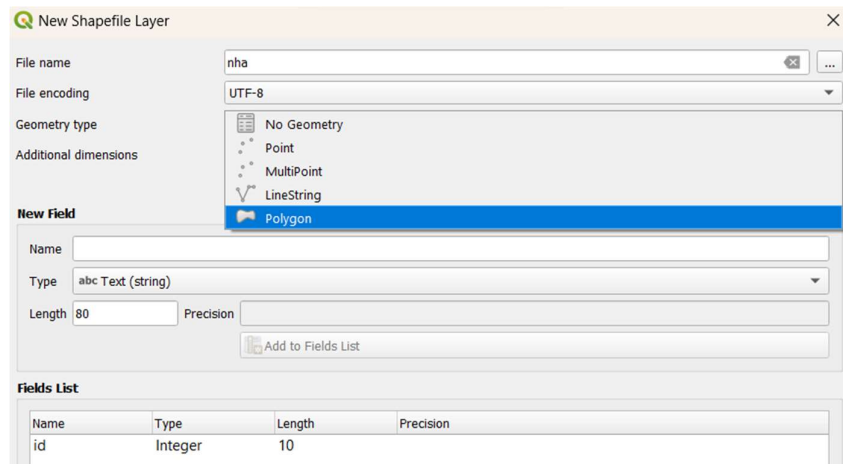
Sử dụng phần mềm QGIS để vector hóa lớp đối tượng trực tiếp trên ảnh theo các bước sau:

Bước 1: Chọn lớp đối tượng để thực hiện số hóa



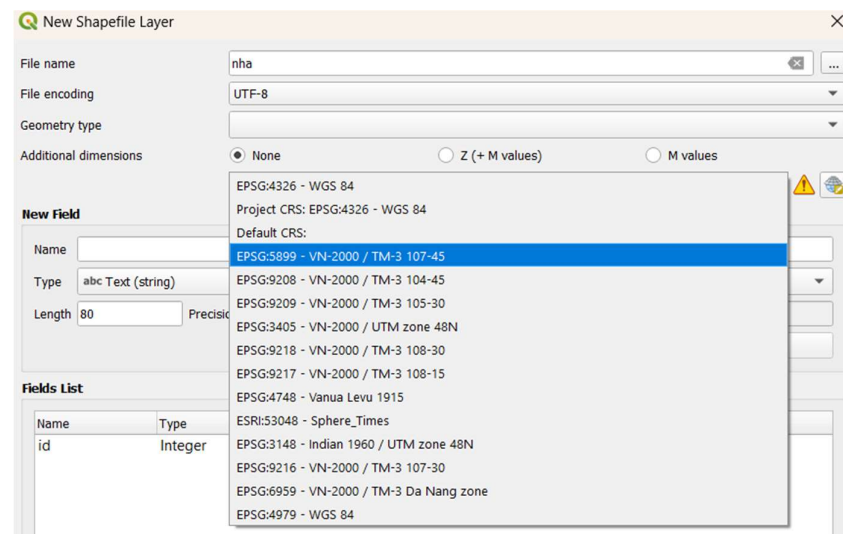
Hình 1-12: Tạo lớp dữ liệu số hóa

Bước 2: Chọn kiểu geometry cho đối tượng cần số hóa



Hình 1-13: Chọn kiểu topology cho đối tượng cần số hóa

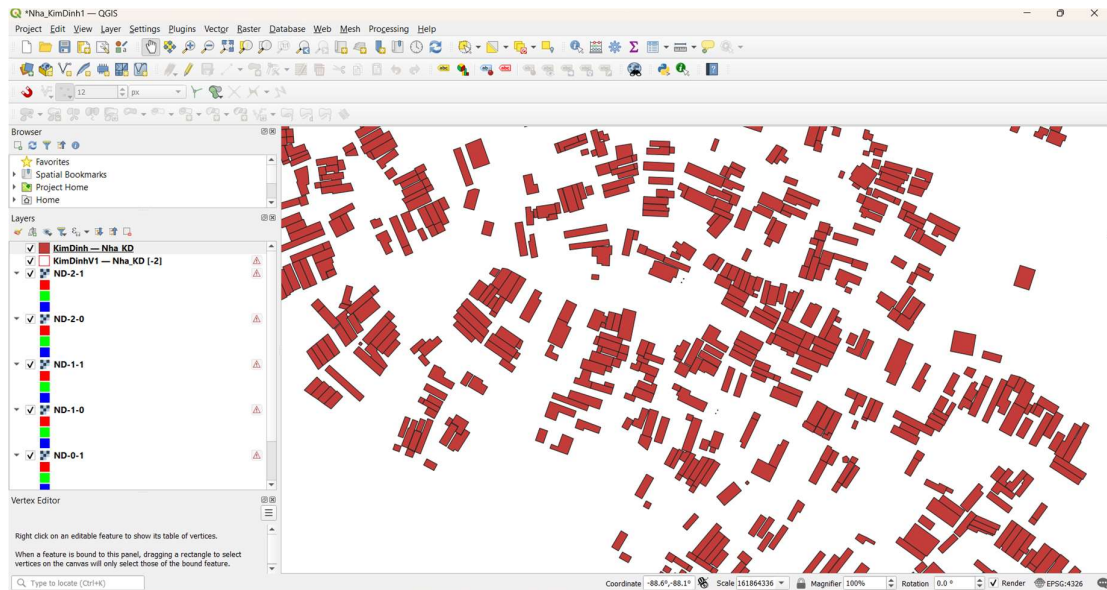
Bước 3: Chọn hệ quy chiếu phù hợp với hệ quy chiếu của khu vực khảo sát



Hình 1-14: Chọn hệ quy chiếu

Bước 4: Thực hiện số hóa lớp đối tượng Nhà

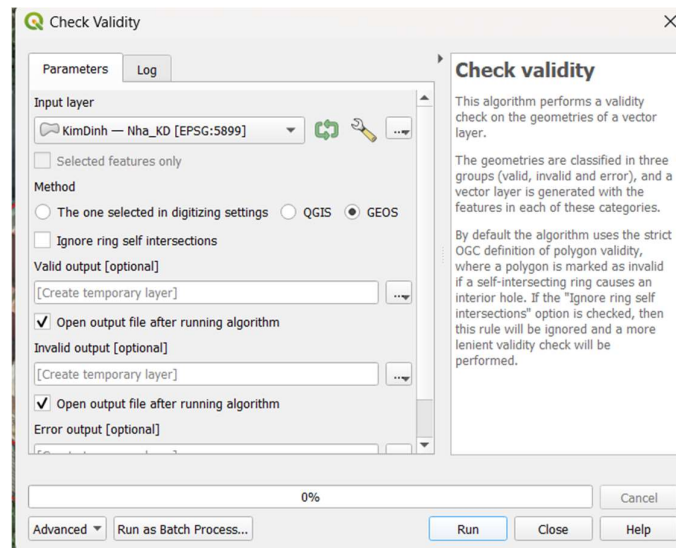
Sau khi đã số hóa các đối tượng (polygon) trên ảnh trực giao như các bước nêu ở bên trên sẽ được các đối tượng như Hình 1-15.



Hình 1-15: Polygon của đối tượng nhà sau khi số hóa trên ảnh UAV

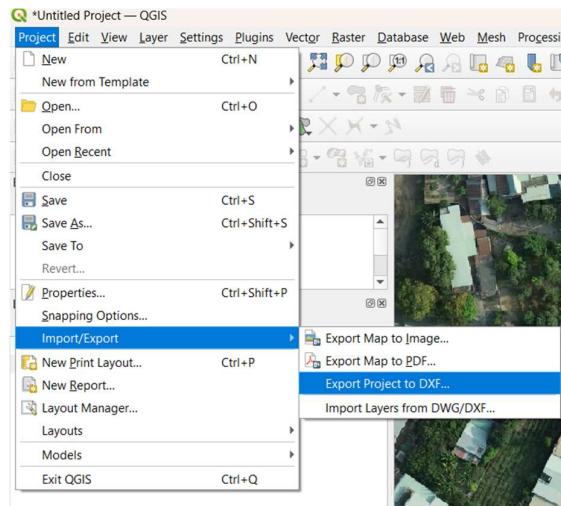
Bước 5: Tiến hành kiểm tra topology

Bước này được thực hiện nhằm kiểm tra các topology bị lỗi trong quá trình số hóa ví dụ bất quá hoặc chưa tới điểm.



Hình 1-16: Kiểm tra topology của đối tượng đã số hóa

Ngoài ra cũng có thể chuyển sang dạng shape file lớp đối tượng đã được số hóa bằng các phần mềm khác như AutoCad hoặc Microstation. Trường hợp này sử dụng công cụ export lớp đối tượng vào phần mềm QGIS như trong



Hình 1-17: Xuất file dữ liệu đã số hóa sang dạng shapefile sử dụng phần mềm QGIS

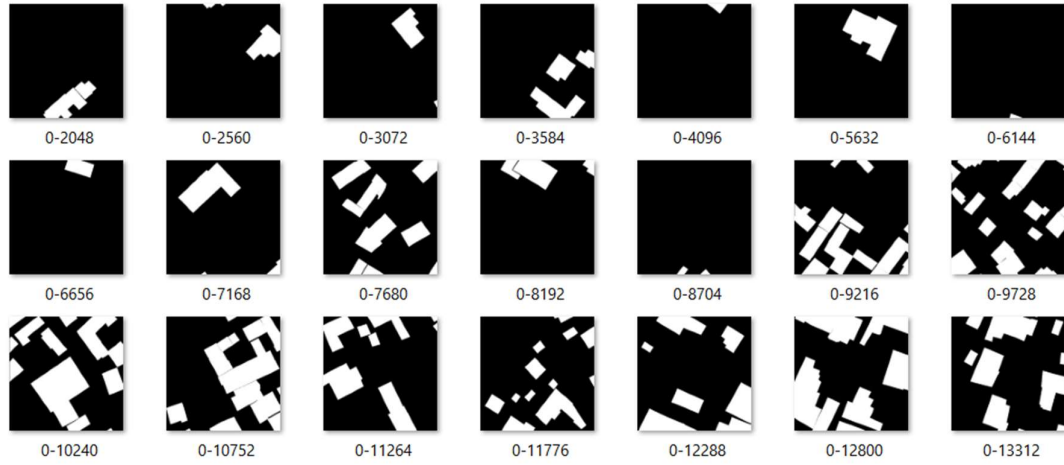
Quá trình tạo mẫu cho mô hình học sâu được thực hiện bằng cách chỉ nhỏ ảnh ortho đầu vào thành nhiều mẫu nhỏ. Vì ảnh UAV sau khi nắn từ nhiều ảnh chụp sẽ có kích thước rất lớn (GiGa Byte như đã thảo luận trong phần ảnh màu) không khả thi cho việc phát triển mô hình học máy sử dụng mạng. Việc đọc toàn bộ hình ảnh vào bộ nhớ và sử dụng nó cho việc phát triển mô hình U-Net là không khả thi. Vì thế, để tăng hiệu suất của quá trình huấn luyện ảnh nắn sẽ được chia thành nhiều mẫu nhỏ gọi là các tiles. Kích thước của tiles phụ thuộc vào cách thiết kế mô hình và thường có kích thước (256 x 256), (512 x 512) hoặc (1024 x 1024) pixel.



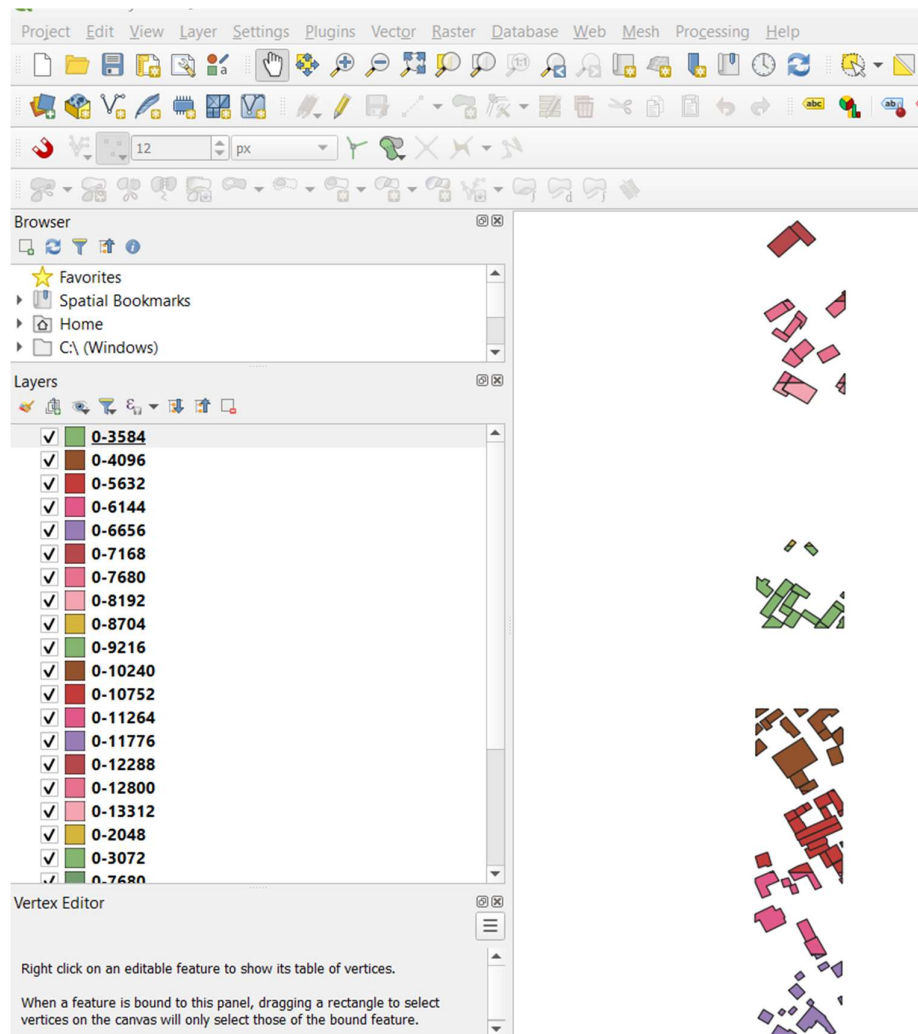
Hình 1-18: Các mẫu tiles được chia nhỏ từ một tấm ảnh UAV ortho lớn

Sau khi đã chia toàn bộ ảnh thành các mẫu nhỏ sẽ thực hiện công việc raster hóa các polygon thành các “mặt nạ - mask” với mức độ xám là 8 bit. Các mask sử dụng để phân biệt đối tượng bởi mã nhị phân. Ví dụ trong Hình 1-19 phần có đối tượng nhà

(tương ứng với màu trắng) được mã là 1 và phần không có đối tượng nhà (tương ứng với màu đen hay là background) có mã là 0.



Hình 1-19: Các mask tương ứng với các tiles



Hình 1-20: Các đối tượng là Nhà được vector hóa tương ứng với các tiles



(a): True ortho- ảnh trực giao

(b): Vector đối tượng nhà



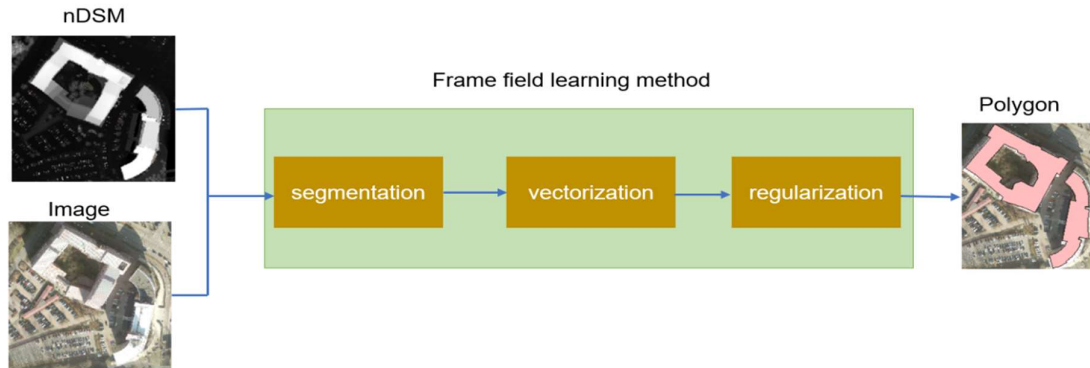
(c): mặt nạ của tòa nhà

Hình 1-21: Một mẫu từ tập dữ liệu trong khu vực nghiên cứu trong đó (a)-ảnh trực giao, (b)- là vector của tòa nhà, và (c) là mask được raster hóa tòa nhà.

1.3 Quy trình trích xuất dữ liệu ảnh sử dụng mạng học sâu

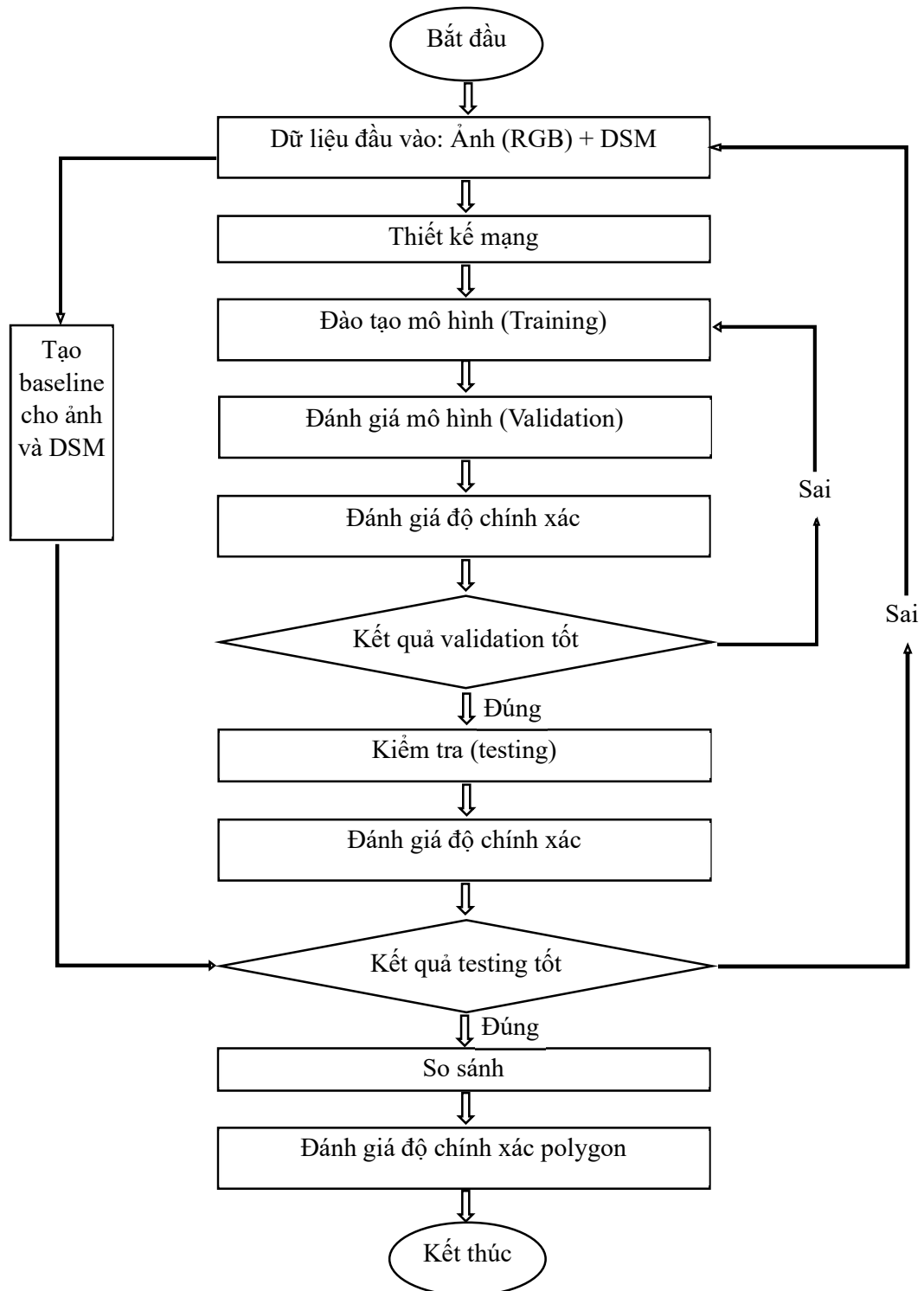
Hầu hết phương pháp học sâu ứng dụng cho trích xuất lớp nhà đều dựa trên việc xây dựng mặt nạ (mask) theo dạng raster, dạng mà không thể sử dụng trực tiếp ở trong dữ liệu GIS. Vì vậy, mô hình học sâu dựa trên phân đoạn đối tượng (semantic

segmentation) tập trung vào việc trích xuất trực tiếp vết tòa nhà dưới dạng các đường polygon. Tuy nhiên, những mô hình này gặp khó khăn về độ chính xác và các hình dạng bất thường của tòa nhà. Hiện nay, phương pháp trích xuất đối tượng nhà dựa trên việc học từ các khung để trích xuất footprints của tòa nhà bởi các vector dạng polygon trực tiếp từ ảnh RGB. Phương pháp FCN được đào tạo để học từ các mặt nạ tòa nhà (mask), đường đồng mức, và các khung bởi phương pháp polygon hóa.



Hình 1-22: Trích xuất đối tượng nhà từ ảnh RGB+ nDSM trên mô hình học sâu với đầu ra là các polygon (Sun 2021)

Ban đầu các mô hình học sâu áp dụng cho ảnh vệ tinh với đầu vào chỉ là các ảnh với ba kênh RGB như được giới thiệu (Girard, Smirnov et al. 2020). Ảnh quang học bị giới hạn bởi thông tin về độ cao nên việc kết hợp với mô hình số bề mặt DSM để cho phép trích xuất đối tượng sẽ góp phần nâng cao độ chính xác. Sơ đồ quy trình trích xuất dữ liệu dựa trên mô hình mạng học sâu từ ảnh (RGB) kết hợp mô hình số độ cao/bề mặt (DSM) được thể hiện như hình Hình 1-22. Baseline được tạo ra đối với cả ảnh (RGB) và DSM phục vụ cho việc so sánh với các polygon được tạo ra từ mô hình. Độ chính xác được đánh giá dựa trên ba mức pixel, đối tượng và polygon (pixel-level, object-level, polygon-level). Thêm vào đó, phân tích độ lệch về số lượng trên số lượng nhà trích xuất bởi mô hình và so sánh với các polygon chuẩn. Đây cũng là một tiêu chí khác để đánh giá chất lượng trích xuất của mô hình nhưng không phải là một tiêu chuẩn. Mạng học sâu được thiết kế dựa trên mạng U-Net được trình bày trong chương 2.



Hình 1-23: Sơ đồ quy trình trích xuất dữ liệu sử dụng mạng học sâu

Chương 2 THỰC NGHIỆM

2.1 Dữ liệu khu vực nghiên cứu

Dữ liệu phục vụ nghiên cứu trong đề án tốt nghiệp này là lớp đối tượng nhà tại một số khu vực dân cư ở cảng biển tại Ngãi Giao, Kim Dinh tỉnh Bà Rịa - Vũng Tàu và khu vực Cam Đức, Cam Ranh tỉnh Khánh Hòa. Lớp đối tượng nhà được số hóa trực tiếp từ ảnh màu UAV (3 kênh phổ) với độ phân giải không gian rất cao từ 5 đến 10cm. Để thực hiện việc đào tạo mô hình học máy, bộ dữ liệu được chia thành ba tập dữ liệu độc lập để phục vụ cho ba mục đích gồm đào tạo (training), đánh giá mô hình (validation), và kiểm chứng thực tế (testing). Sau đây sẽ trình bày cụ thể ba tập dữ liệu này.

2.1.1 Tập dữ liệu huấn luyện mô hình (*Training data set*)

Training data set là một tập dữ liệu dùng để huấn luyện cho mô hình của thuật toán học máy (Machine Learning). Đây thường là một tập dữ liệu có kích thước lớn, được dùng để training trong quá trình huấn luyện học máy. Training data set bao gồm bộ dữ liệu đầu vào (input) và bộ dữ liệu đầu ra (output) tương ứng, trong đó output thường được gọi là nhãn (label hoặc target). Các thuật toán học máy nói chung sẽ tìm cách tối ưu sai số dự đoán trên tập huấn luyện này đến mức “đủ” tốt theo từng bài toán và mục tiêu ứng dụng cụ thể.

Trong nghiên cứu này, bộ dữ liệu phục vụ việc huấn luyện (training) mô hình được tổng hợp từ 5 khu vực khác nhau tại thị trấn Ngãi Giao, Huyện Châu Đức, tỉnh Bà Rịa-Vũng Tàu với 2829 mẫu. Các file được sử dụng gồm: *Ngãi giao khu vực 1, Ngãi giao khu vực 2, Ngãi giao khu vực 3, Ngãi giao khu vực 4, Ngãi giao khu vực 5*.

2.1.2 Tập dữ liệu đánh giá hiệu năng mô hình học máy trong quá trình huấn luyện (*Validation data set*)

Sau khi xây dựng một mô hình học máy và huấn luyện nó trên một tập dữ liệu huấn luyện (training data set), bước tiếp theo là đánh giá hiệu năng của mô hình trên tập

dữ liệu mới độc lập với tập dữ liệu đã phục vụ cho quá trình huấn luyện ở trên. Giống như training data set, validation data set cũng bao gồm các cặp giá trị input và output tương ứng. Tuy nhiên, input & output của tập validation data set không dùng để huấn luyện trong quá trình học máy mà chỉ sử dụng để kiểm tra độ chính xác của mô hình học máy trong quá trình huấn luyện. Nói cách khác, validation data set là tập các giá trị input đi kèm với giá trị output và được dùng để kiểm thử độ chính xác của mô hình học máy trong quá trình huấn luyện (training).

Trong nghiên cứu này, tập dữ liệu phục vụ cho việc đánh giá hiệu quả của mô hình trong quá trình huấn luyện (validation) mô hình được thu thập từ khu vực phường Kim Dinh tỉnh Bà Rịa-Vũng Tàu với 457 mẫu.

2.1.3 Tập dữ liệu đánh giá hiệu năng mô hình học máy sau quá trình huấn luyện (Testing data set)

Nói một cách dễ hiểu, Testing data set là tập dữ liệu dùng để kiểm tra mô hình sau khi máy tính đã học xong. Một mô hình học máy sau khi được huấn luyện sẽ cần phải được kiểm chứng xem nó có đạt hiệu quả không thông qua việc đánh giá độ chính xác. Nếu độ chính xác của mô hình đáp ứng được yêu cầu thì mô hình được coi là đạt yêu cầu. Ngược lại nếu độ chính xác không đạt yêu cầu thì cần phải đào tạo lại mô hình hay xác định lại các siêu tham số (hyperparameters) của mô hình. Để kiểm tra, đánh giá độ chính xác của mô hình đối với phương pháp học máy, tập kiểm tra (testing data set) được sử dụng. Khác với training data set, testing data set chỉ gồm các giá trị input mà không có giá trị output. Nói một cách ngắn gọn, testing data set là tập các giá trị input và được dùng để kiểm thử độ chính xác của những mô hình học máy sau khi được huấn luyện. Vậy sự khác nhau giữa công dụng testing data set và validation data set là gì? Testing được dùng để kiểm mô hình “sau” quá trình huấn luyện, còn validation set được sử dụng để kiểm mô hình “trong” quá trình huấn luyện.

Trong nghiên cứu này, tập dữ liệu phục vụ cho việc đánh giá hiệu của mô hình sau quá trình huấn luyện (testing) mô hình được thu thập từ khu vực Cam Đức, huyện Cam Lâm và thành phố Cam Ranh tỉnh Khánh Hòa với 538 mẫu.

2.3 Phát triển mô hình Unet để phân loại đối tượng nhà

Kiến trúc của mạng Unet giống như hình chữ ‘U’ đặc điểm này giải thích cho tên gọi của nó. Kiến trúc này gồm 3 phần: contraction (encoder), bottleneck và expansion (decoder). Phần contraction được tạo bởi nhiều khối (block), mỗi khối gồm hai lớp CNN có kernel size là 3×3 , theo sau bởi một lớp 2×2 max pooling. Số lượng kernel hay feature map tăng gấp đôi sau mỗi khối, do đó kiến trúc này có thể học những đặc trưng phức tạp một cách hiệu quả. Lớp ở vị trí thấp nhất (bottleneck) làm trung gian giữa các lớp ở phần contraction và phần expansion. Nó dùng 2 lớp 3×3 CNN, theo sau bởi 1 với 2×2 up convolution.

Trái tim của kiến trúc này nằm ở phần expansion (decoder). Tương tự như contraction, nó cũng bao gồm nhiều khối, mỗi khối gồm hai lớp 3×3 CNN theo sau bởi một lớp upsampling. Sau mỗi khối số lượng feature map được giảm đi một nửa để đảm bảo tính đối xứng. Tuy nhiên đầu vào của mỗi khối expansion sẽ được nối thêm feature map của khối contraction tương ứng. Sự kết nối này sẽ đảm bảo các được trung được học trong quá trình contracting ảnh sẽ được dùng để tái tạo nó. Số lượng khối expansion bằng với số lượng khối contraction. Sau cùng, kết quả sẽ được đưa qua một lớp 3×3 CNN khác với số feature map bằng với số vùng mong muốn. Kiến trúc của mạng Unet cơ bản và mạng Unet được đề xuất thể hiện trong.

Trong quá trình phát triển mô hình Unet sử dụng thư viện Pytorch, nghiên cứu đã kiểm tra độ chính xác của cả dữ liệu đào tạo và đảm bảo rằng mô hình không bị quá khít (overfitting) hoặc không khớp (underfitting) với mẫu đầu vào. Mô hình Unet tốt nhất sẽ được chọn dựa trên khả năng dự đoán chính xác đối tượng lớp nhà và phù hợp với các nhãn được chỉ định từ dữ liệu huấn luyện trong tập dữ liệu thô. Nghiên cứu đã triển khai mô hình Unet bằng thư viện Pytorch, sử dụng API mô hình Segmentation trong khuôn khổ của Pytorch. Nghiên cứu đã lựa chọn sử dụng Pytorch thay vì Tensorflow trong quá trình phát triển mô hình Unet vì một số lý do cụ thể sau:

Đầu tiên, Pytorch cung cấp một cách tiếp cận đơn giản và trực quan hơn trong quá trình xây dựng và đào tạo mô hình. Cú pháp của Pytorch giống với việc lập trình

Python thông thường, giúp làm giảm độ phức tạp và tăng tính hiệu quả của quá trình phát triển.

Thứ hai, Pytorch có một cộng đồng lớn và tích hợp các tính năng mới và cải tiến một cách linh hoạt và nhanh chóng. Điều này giúp cho việc tiếp cận và tích hợp những phát triển mới trong lĩnh vực học máy và thị giác máy tính một cách dễ dàng hơn.

Thứ ba, Pytorch thường cho phép triển khai mô hình một cách linh hoạt hơn, đặc biệt là khi cần tích hợp mô hình vào các ứng dụng thực tế. Pytorch có cộng đồng hỗ trợ mạnh mẽ cho việc triển khai trên nhiều nền tảng và thiết bị khác nhau.

Tóm lại, sự linh hoạt, trực quan và tích hợp nhanh chóng của Pytorch đã làm cho nó trở thành lựa chọn ưa thích trong dự án phát triển mô hình Unet của nhóm nghiên cứu. Mô hình Unet được thiết kế để phân đoạn hình ảnh và được xây dựng dựa trên Pytorch thay vì Tensorflow. Chu kỳ đào tạo của mô hình Unet được giới hạn ở 200 vòng (epoch), tuy nhiên, nếu hệ số chính xác trên tập dữ liệu thử nghiệm hội tụ, chu kỳ có thể bị dừng lại nếu tất cả các giá trị chính xác không thay đổi sau 20 chu kỳ.

2.4 Kết quả dự đoán (trích xuất) lớp nhà từ mô hình

Như đã trình bày trong Chương 1 về cách đánh giá mô hình, theo đó độ chính xác của mô hình được đánh giá thông qua hai bước gồm đánh giá trong quá trình đào tạo mô hình (validation) và đánh giá sau quá trình đào tạo mô hình (testing). Trong đề tài này ba bộ dữ liệu sử dụng để đào tạo (training) mô hình, validation và testing được sử dụng ở ba khu vực có đặc điểm đối tượng cần trích xuất (lớp nhà) tương tự nhau. Như đã nêu ở phần 4.1, mô hình được học từ 2829 mẫu ảnh có kích thước 512x512 pixel tại Ngãi Giao, Huyện Châu Đức, tỉnh Bà Rịa-Vũng Tàu.

Từ các siêu tham số (hyperparameters) xác định được trong quá trình huấn luyện, tiến hành đánh giá độ chính xác mô hình (validation) thông qua tập mẫu đã được chuẩn bị. Tập validation cũng là các mẫu ảnh có kích cỡ 512x512, trong thực nghiệm này, tập validation gồm 457 mẫu ảnh tại phường Kim Dinh tỉnh Bà Rịa-Vũng Tàu. Lưu ý trong chương trình máy tính hai bước training và validation được thực hiện liên tiếp nhau trong cùng một chương trình máy tính. Sau đây sẽ trình bày kết quả đánh giá độ chính xác của hai quá trình training và validation.

Bảng 4-3 trình bày kết quả đánh giá độ chính xác của training và validation khi sử dụng ba loại hàm loss function gồm CeLoss (cross entropy), DiceLoss, và FocalLoss.

- Đối với training mô hình:

Độ chính xác (training accuracy) đào tạo mô hình của hàm CeLoss đạt giá trị tốt hơn (0.99) khi so với trường hợp mô hình sử dụng hàm DiceLoss và FocalLoss là 0.96 và 0.97, tương ứng. Các giá trị của hàm mất mát `train_loss_epoch` và `train_loss_step` cũng có xu hướng tương tự như giá trị độ chính xác training. Giá trị `Train_loss_epoch` và `train_loss_step` khi sử dụng Celoss là 0.003 và 0.004 so với khi sử dụng hàm DiceLoss là 0.069 và 0.056 và FocalLoss là 0.022 và 0.019, tương ứng.

- Đối với validation mô hình:

Mô hình được đánh giá trong quá trình đào tạo validation accuracy có giá trị hợp lý so với giá trị khi đào tạo training accuracy. Giá trị của validation accuracy khi áp dụng cả ba hàm CeLoss, DiceLoss và FocalLoss đều nhỏ hơn so với giá trị training accuracy. Trong đó, validation accuracy khi sử dụng hàm mất mát CeLoss cũng có độ chính xác cao hơn chút ít so với hai hàm còn lại. Giá trị đạt được là 0.942 (CeLoss) so với 0.926 (DiceLoss) và 0.937 (FocalLoss). Tuy nhiên giá trị mất mát `val_loss_epoch` khi sử dụng CeLoss lại cao bất thường 0.59 mặc dù số epoch sử dụng lên tới 60 trong khi sử dụng hai là mất mát còn lại chỉ mất 20 epoch cho mô hình hội tụ (xem chi tiết trong các. Do đó cần phải xem xét lại mô hình trong trường hợp này. Hai yếu tố nên phải xem xét để đào tạo mô hình đó là số lượng mẫu và độ chính xác tạo mẫu. Bởi vì hiện tại mô hình mới chỉ sử dụng dưới 3000 mẫu nên có thể chưa đảm bảo sự đa dạng về mẫu. Các mẫu cũng cần được xem xét lại trong quá trình vector hóa đối tượng nhà ví dụ trường hợp tạo đường bao gộp nhiều nhà liền kề thành một khu nhà hoặc gộp hình ảnh của mái hiên, nhà bếp... với nhà chính thành một đối tượng nhà. Bảng 0-1: So sánh kết quả đánh giá mô hình với số lượng epoch là 40 và 60 Ngoài chỉ tiêu đánh giá (metrics) theo độ chính xác tổng thể (accuracy) thì còn có các thước đo khác đáng tin cậy hơn đối với bài toán phân đoạn hình ảnh (segmentation) như chỉ số IoU (tỉ số vùng giao và vùng hợp) và độ đo Dice như đã trình bày trong Chương 2. Phần này sẽ tổng hợp kết quả đánh giá mô hình thông qua chỉ số IoU với hai lớp nhà (building) và lớp background.

Các chỉ số độ chính xác accuracy và theo IoU của cả lớp building và background đều tăng theo số lượng vòng lặp epoch tăng lên từ 40 đến 60 và mô hình có xu hướng hội tụ. Độ chính xác tổng thể và theo IoU của lớp nhà đạt mức 0.99 và 0.97 tại epoch=60 tương ứng với hai bước training và validation. Bên cạnh đó độ chính xác của lớp background cũng đạt được 0.88 và 0.87 với bước training và validation tại epoch =60. Giá trị trung bình mIoU của hai lớp building và background cũng xác định được với giá trị là 0.927 tại epoch =60 cho cả hai bước trên. Giá trị hàm mất mát của bước training giảm từ 0.0057 giảm còn 0.0099 khi số vòng lặp tăng lên từ 40 đến 60. Giá trị này là phù hợp và thể hiện rằng mô hình hội tụ với các tham số đã chọn. Tuy nhiên ở bước đánh giá validation thì giá trị hàm mất mát val_loss_epoch lại tăng từ 0.244 đến 0.484 với số vòng lặp tăng lên từ 40 lên 60 (có thể xem sự thay đổi này trong ...). Kết quả này cho thấy mô hình ở bước đánh giá validation vẫn chưa đạt yêu cầu về đào tạo mô hình, cho nên cần phải tiếp tục điều chỉnh lại mẫu (số lượng mẫu tăng lên hoặc cách tạo mẫu) hoặc thay đổi các tham số của mô hình... Tuy nhiên trong giới hạn của đề tài các vấn đề nâng cao độ chính xác mô hình chưa được trình bày ở đây.

- Đối với đánh giá mô hình sau khi đào tạo mô hình (testing)

Đánh giá mô hình sau khi đã đào tạo là quá trình testing. Để thực hiện bước này đề tài đã sử dụng 538 mẫu ảnh có kích thước 512x512 pixel tại khu vực Cam Đức, tỉnh Khánh Hòa. Kết quả dự đoán đối tượng nhà bằng mô hình Unet đã được đào tạo được thể hiện trên Hình 0-3 và so sánh với đối tượng thực trên Hình 0-4.



Hình 0-1: Ảnh UAV sau khi nắm khu vực kiểm tra testing



Hình 0-2: Lớp nhà được số hóa (vector) sử dụng làm đối tượng thực để testing



Hình 0-3: Kết quả dự đoán từ mô hình Unet



Hình 0-4: So sánh kết quả dự đoán (đường màu đỏ) và đối tượng thực (màu vàng)

Quá trình số hóa tự động đối tượng nhà bằng mô hình Unet đã đạt được một số kết quả nhất định. Mô hình đã phát hiện được đối tượng nhà thể hiện bằng đường màu đỏ (dự đoán) đã nằm trên các đối tượng nhà trên hình ảnh UAV và phù hợp với mẫu (màu vàng) đã số hóa. Tuy nhiên tỉ lệ đường màu đỏ và vàng trùng khớp nhau vẫn còn thấp điều này cho thấy độ chính xác dự đoán từ mô hình còn thấp. Đặc biệt có một số vị trí có mẫu nhà mà mô hình chưa phát hiện ra hoặc vị trí không có nhà mô hình lại dự đoán sai. Điều này cho thấy cần phải điều chỉnh lại mô hình và học thêm với số lượng mẫu lớn hơn bởi vì mô hình đang chỉ sử dụng đào tạo với dưới 3 nghìn mẫu.

KẾT LUẬN VÀ KIẾN NGHỊ

Quy trình trích xuất dữ liệu từ ảnh UAV được đề xuất trong đề án là một kết quả quan trọng nhằm bổ sung thêm cơ sở khoa học và có ý nghĩa thực tiễn trong thông tư 07/2021 của Bộ TN&MT về việc cho phép thành lập bản đồ địa hình tỉ lệ lớn và cơ sở dữ liệu từ thiết bị bay không người lái UAV. Quy trình đã nêu đầy đủ các bước từ thu thập dữ liệu bay UAV, tạo mẫu cho quá trình học sâu, đánh giá kết quả từ bộ dữ liệu đào tạo và cuối cùng là trích xuất dữ liệu. Thông qua thực nghiệm, kết quả trích xuất dữ liệu đã bước đầu đáp ứng được mục tiêu đề ra về phân loại đối tượng nhà. Điều này minh chứng cho quy trình đề ra là hợp lý và có tính thực tiễn cao. Tuy nhiên vẫn cần chú ý đến công tác xây dựng bộ mẫu và cải thiện các mô hình học sâu trong công tác đào tạo mô hình để mang lại kết quả với độ chính xác cao hơn.

TÀI LIỆU THAM KHẢO

- "Thông tư số 07/2021/tt-btnmt quy định kỹ thuật thu nhận và xử lý dữ liệu ảnh số từ tàu bay không người lái phục vụ xây dựng, cập nhật cơ sở dữ liệu nền địa lý quốc gia tỷ lệ 1:2.000, 1:5.000 và thành lập bản đồ địa hình tỷ lệ 1:500, 1:1.000."
- Cường, N. S., T. X. Trường, T. H. Hạnh and Đ. N. J. T. c. K. h. K. t. M.-Đ. c. Vũ, 60 (2019). "Nâng cao chất lượng xây dựng mô hình 3d bằng kết hợp công nghệ bay chụp uav và quét laser mặt đất." 10.
- Girard, N., D. Smirnov, J. Solomon and Y. J. a. p. a. Tarabalka (2020). "Polygonal building segmentation by frame field learning."
- Lê, V. C., T. T. H. Lê and X. C. Cao (2020). "Nghiên cứu lựa chọn vị trí cất cánh cho thiết bị bay không người lái tích hợp gnss động phục vụ đo vẽ thành lập bản đồ địa hình tỷ lệ lớn cho các mỏ lộ thiên= research on optimal takeoff positions of uav integrated gnss-rtk in producing large scale topological maps for open-pit mines."
- Long, N. Q. and C. X. J. T. c. c. n. m. Cường (2019). "Ứng dụng máy bay không người lái (uav) để xây dựng mô hình số bề mặt và bản đồ mỏ lộ thiên khai thác vật liệu xây dựng." 1: 21-29.
- Long, N. Q., R. Goyal, L. K. Bui, C. X. Cuong, L. V. Canh, N. Q. Minh and X. N. J. A. o. M. S. Bui (2021). "Optimal choice of the number of ground control points for developing precise dsm using light-weight uav in small and medium-sized open-pit mine." 66(3).
- Sun, X. (2021). Deep learning-based building extraction using aerial images and digital surface models, University of Twente.
- Sỹ, M. V., B. N. Quý and P. V. J. T. c. K. h. Đ. đ. v. B. đ. Hiệp (2017). "Nghiên cứu sử dụng dữ liệu ảnh máy bay không người lái (uav) trong thành lập bản đồ địa hình tỷ lệ lớn." (33): 49-57.