

ĐỀ CƯƠNG NGHIÊN CỨU KHOA HỌC

Lĩnh vực: Phần mềm hệ thống

PHẦN MỀM PHÁT NGUYỆ: PHÁT HIỆN NGƯỜI KHÔNG ĐỘI MŨ BẢO HIỂM

Ứng dụng thuật toán YOLOv8 để phát triển một công cụ quét thông minh chuyên phát hiện các trường hợp vi phạm không đội mũ bảo hiểm khi tham gia giao thông.



Người thực hiện: Đồng Nhật Diễm My 11B10
Hoàng Nguyễn Ngọc Khoa 11B01
Giáo viên hướng dẫn: cô Nguyễn Thị Thà

Tháng 8/2024 - TP. Hồ Chí Minh

MỤC LỤC

MỤC LỤC.....	1
TÓM LƯỢC ĐỀ TÀI.....	2
Tổng quan.....	2
Tính mới.....	2
Từ khóa:.....	2
1. Phần mở đầu.....	2
2. Thông tin nghiên cứu.....	3
2.1. Giới thiệu đề tài.....	3
Thuật toán YOLO là gì?.....	4
2.2. Giả thuyết khoa học và tiêu chí đánh giá.....	4
2.3 Phương pháp nghiên cứu.....	4
3. Tiến hành nghiên cứu.....	6
3.1. Chuẩn bị Dataset.....	6
3.2. Lập trình.....	6
3.3. Huấn luyện mô hình - Số liệu.....	7
3.3.1. Biểu đồ huấn luyện tổng quan.....	7
3.3.2. Biểu đồ loss.....	8
3.3.3. Biểu đồ metrics.....	9
3.3.4. Biểu đồ phân tán F1 scores.....	10
3.4. Kết quả.....	11
3.5. Thảo luận.....	12
3.5.1. Điểm tốt.....	12
3.5.2.Điểm chưa tốt.....	13
4. Kết luận của đề tài.....	13
5. Khả năng phát triển đề tài.....	13
TÀI LIỆU THAM KHẢO.....	14

TÓM LƯỢC ĐỀ TÀI

Tổng quan

Ngày nay, ứng dụng về Thị Giác Máy Tính (TGMT) trong Trí Tuệ Nhân Tạo (TTNT) đang ngày càng trở nên phổ biến hơn do sự tiến bộ vượt bậc của công nghệ. Đặc biệt gần đây, TGMT đang bắt đầu có sự chuyển hướng vào việc thiết lập một hệ thống Giao Thông Thông Minh (GTTM) và an toàn hơn. Trong nghiên cứu này, chúng tôi sẽ tập trung xây dựng một công cụ tương tự dành cho các quốc gia tiêu thụ lượng lớn xe máy, thông qua chức năng chính là phát hiện vi phạm người không đội mũ bảo hiểm.

Tính mới

Đề tài sử dụng các công nghệ hiện đại như AI, machine learning (thuật toán YOLO), và xử lý ảnh (OpenCV) để phát hiện và nhận diện người không đội mũ bảo hiểm, một ứng dụng còn mới mẻ và đang có tiềm năng phát triển mạnh trong giao thông Việt Nam.

Với chức năng phát hiện vi phạm tự động: Giải pháp này giúp tự động hóa quá trình kiểm soát và đảm bảo an ninh trật tự giao thông mà không cần sự can thiệp trực tiếp từ con người. Trong tương lai còn có thể thay thế phương pháp thủ công truyền thống, tiết kiệm triệt để thời gian và nguồn lực.

Từ khóa:

phát hiện đối tượng, phát hiện mũ bảo hiểm, phạt nguội, mũ bảo hiểm, thị giác máy tính

1. Phần mở đầu

Trong khi phong trào đô thị hóa ở Việt Nam có sự phát triển nhanh chóng thì mối lo ngại về tai nạn giao thông ngày càng tăng khi số lượng xe ra vào mỗi lúc ngày càng nhiều. Theo nhiều báo cáo, những năm trở lại đây, số lượng xe máy luôn chiếm hơn 80% trên tổng số phương tiện được đăng ký trên cả nước [1-3]. Tính riêng từ năm 2018-2022 đã có trung bình khoảng 3 triệu xe máy tăng lên mỗi năm [4], cho thấy sự tiêu thụ xe máy ồ ạt như một nét văn hóa đặc trưng của người Việt.

Nhìn chung, Việt Nam là nước đang phát triển với đặc thù giao thông đường sá chật hẹp, lượng xe cộ ra vào càng nhiều đã kéo theo hệ lụy với những tai nạn giao thông (TNGT) vô cùng thương tiếc. Theo các nghiên cứu, nguyên nhân chính của các vụ tai nạn giao thông bởi xe máy là do hành vi của người tham gia giao thông, thiếu hiểu biết về luật lệ và văn hóa ứng xử khi tham gia giao thông [5]. Một trong những quy tắc “vàng” thường bị bỏ qua đó chính là đội mũ bảo hiểm (MBH). Điều này được chứng minh trong một đánh giá toàn cầu

(Abbas và nnk. 2012), kết quả đã cho thấy rằng nguyên nhân chủ yếu nhất ảnh hưởng đến tỷ lệ tử vong của người đi xe máy trong các vụ TNGT chính là không đội MBH. Đánh giá cũng cho thấy MBH làm giảm nguy cơ tử vong đáng kể khi xảy ra va chạm.

Khi không được bảo vệ bởi MBH một cách kỹ lưỡng, sẽ làm tăng nguy cơ các chấn thương về đầu và sọ não, chấn thương cổ... Chấn thương càng nghiêm trọng, nạn nhân càng phải đối mặt với việc điều trị và phục hồi sau tai nạn phức tạp và tốn kém hơn [7]. Kết luận, chỉ việc không chấp hành việc đội MBH khi tham gia giao thông cũng có thể gây nên hậu quả đáng lo ngại, gây ảnh hưởng đến đời sống cá nhân, gia đình và xã hội. Thực tế nói chung, đại đa số nhân dân nói chung biết rõ đây là hành vi gây mất ATGT nhưng dường như vẫn cố tình vi phạm, theo thời gian mà trở thành thói quen [8]. Sự thờ ơ này cho thấy việc nâng cao ý thức của người dân vẫn chưa đạt hiệu quả như mong đợi.

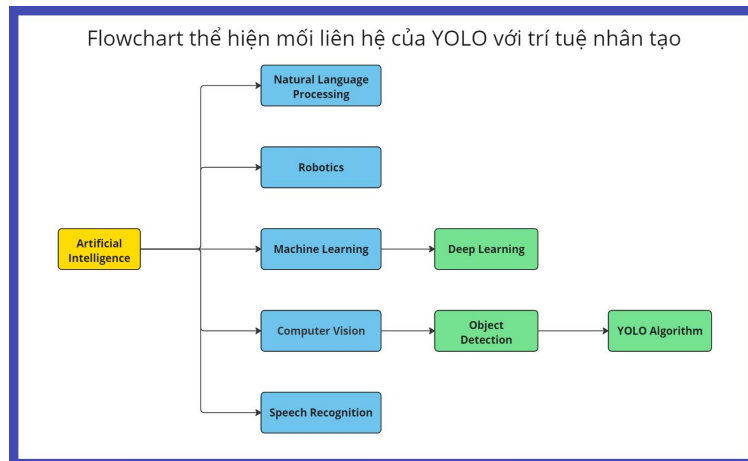
“Đội mũ bảo hiểm không phải đối phó xử phạt với công an mà chính là bảo vệ sức khỏe bản thân mỗi người” (Hạnh, 2019).

2. Thông tin nghiên cứu

2.1. Giới thiệu đề tài

Để khắc phục tình trạng này, cần có một biện pháp vừa khả quan, vừa hiện đại, đảm bảo tăng cường hiệu quả ý thức chấp hành luật của người đi xe máy. Điều này đòi hỏi một hệ thống có độ chính xác, độ tin cậy và độ thông minh đủ cao để đáp ứng với sự thay đổi không ngừng của xã hội. Nhờ Trí Tuệ Nhân Tạo (Artificial Intelligence), chúng ta hoàn toàn có thể phát triển các hệ thống đưa quyết định và dự đoán dựa trên cơ sở dữ liệu [9]. Trong nghiên cứu này, chúng tôi đã tận dụng thuật toán YOLO (You Only Look Once) trong lĩnh vực này để đem đến một sản phẩm thú vị:

| Phần mềm phạt nguội: phát hiện người không đội mũ bảo hiểm”



Hình 1) Flowchart thể hiện mối liên hệ của YOLO với trí tuệ nhân tạo

Thuật toán YOLO là gì?

YOLO (You Only Look Once) là thuật toán phát hiện đối tượng theo thời gian thực được phát triển vào năm 2015 bởi Joseph Redmon và Ali Farhadi [10]. Nó sử dụng mô hình mạng nơ-ron tích chập (CNN) để dự đoán các bounding boxes¹ và class probabilities² trong dữ liệu đầu vào [11].

Tính tới thời điểm hiện tại (tháng 8/2024), YOLO vẫn đang được xem là thuật toán hiện đại, chính xác và nhanh nhất trong lĩnh vực phát hiện đối tượng ở thời gian thực. Đã có nhiều phiên bản được phát triển kể từ khi thuật toán mới được ra mắt, những phiên bản mới nhất có thể kể đến như YOLOv8 (tháng 1/2023), YOLOv9 (tháng 2/2024) và YOLOv10 (tháng 6/2024). Vì để tránh những rủi ro của phiên bản mới phát triển, dự án nghiên cứu này tập trung xây dựng kỹ thuật dựa trên YOLOv8.

2.2. Giả thuyết khoa học và tiêu chí đánh giá

Nghiên cứu này đưa ra giả thuyết rằng phần mềm phát hiện mũ bảo hiểm của chúng tôi sẽ xác định người đi xe máy đội MBH và không đội MBH với tỷ lệ chính xác ít nhất 80% trong điều kiện thường.

Xác suất xác định đúng đối tượng vi phạm:

- + Trong điều kiện thường: $\geq 80\%$.
- + Trong điều kiện thiếu sáng (ban đêm, dưới chân cầu...): 60~80%
- + Trong trời mưa: 50~60%

¹ Bounding box (hộp giới hạn): là một đường viền hình chữ nhật được vẽ xung quanh một đối tượng để định vị đối tượng đó.

² Class probabilities là các xác suất mà mô hình dự đoán cho mỗi lớp, thể hiện mức độ tin cậy rằng một mẫu thuộc về một lớp cụ thể.

2.3 Phương pháp nghiên cứu

Trong quá trình thực hiện, chúng tôi sử dụng 4 phương pháp nghiên cứu : Văn bản tài liệu, Xử lý số liệu, Thực nghiệm và Trường hợp để định hình nền móng cho nghiên cứu:

1) Văn bản tài liệu***	Đây là một trong những phương pháp đóng vai trò quan trọng trong dự án nghiên cứu này. Sở dĩ dữ liệu được cập nhật liên tục, phương pháp này giúp nắm bắt thông tin một cách sơ lược trong quá trình làm việc. Về cơ bản, khi bắt tay vào làm một giai đoạn mới, chúng tôi sẽ chọn tìm đến các nguồn tài nguyên trên internet bằng tiếng Anh, ví dụ bao gồm một số diễn đàn Q&A, Github, Tensorflow, Roboflow, Youtube,... Có thể dùng để giải đáp thắc mắc, xem code mẫu hoặc fix bug...
2) Xử lý số liệu***	Đây cũng là một trong những phương pháp không thể thiếu của dự án. Xử lý số liệu đặc biệt quan trọng từ những bước đầu, để đảm bảo kết quả đưa ra được tối ưu và đạt chuẩn yêu cầu nhất có thể. Chúng tôi sử dụng phương pháp này trong khi lọc, đếm và quản lý số liệu. Đặc biệt là khi làm việc với dataset, ngay từ những bước đầu thu thập hình ảnh cho tới khi tiến hành training, phương pháp này được ứng dụng mạnh mẽ để đánh giá hiệu suất của mô hình, từ đó thay đổi lại dữ liệu và tham số sao cho phù hợp. Sau khi train xong dữ liệu, PPNC này tiếp tục được ứng dụng để phân tích hiệu suất của mô hình qua các biểu đồ và số liệu.
3) Thực nghiệm	- Thử nghiệm chương trình bằng dữ liệu mà mô hình chưa từng gặp (unseen data) qua webcam, video, ảnh trích xuất trên mạng và đánh giá tỉ lệ chuẩn xác của nó sau khi quét.
4) Trường hợp	- Phân loại trường hợp người đi xe đội nón bảo hiểm và không đội nón. - Phân loại môi trường để lập độ tin cậy tiêu chuẩn (dưới chân cầu, ban đêm, trời mưa...) - Loại bỏ các trường hợp người đi bộ.

➡ Với 4 phương pháp này, chúng tôi sẽ tiếp cận theo hướng **chỉ dự đoán trên người đang chạy xe máy**.

3. Tiến hành nghiên cứu

3.1. Chuẩn bị Dataset

Bước đầu tiên và cũng là một trong những bước quan trọng nhất đó là chuẩn bị dataset. Trong dự án này, chúng tôi tạo ra 2 dataset, dataset phát hiện xe máy, (gồm 1143 ảnh) và dataset mũ bảo hiểm (gồm 5484 ảnh).

Riêng dataset mũ bảo hiểm, tuy có các hình được chụp thực tế nhưng phần lớn là thu thập từ trên mạng. Để có thể xác định được một loại vật thể, việc phân loại lớp (Class) cho vật thể đó và đặt nhãn (Label) cho nó là điều thiết yếu.

Chúng tôi đã phân ra 2 class với label là Helmet (có đội mũ bảo hiểm) và NonHelmet (không đội mũ bảo hiểm).

Đối với class Helmet, chúng tôi thu thập những hình ảnh có chứa MBH. Ngoài việc lấy ảnh từ trên mạng, chúng tôi cũng đảm bảo thu thập hình ảnh từ thực tế để mô hình học được nhiều nhất, chẳng hạn như MBH Trần Phú. Còn đối với class Nonhelmet sẽ có phần đặc biệt hơn, class này sẽ được xác định qua các đặc điểm đặc trưng như đầu trần, kiểu tóc, màu tóc,... được chụp ở nhiều góc độ để đảm bảo dữ liệu được đa dạng nhất.

Dataset của chúng tôi được định dạng theo cấu trúc YOLOv8³, và chia ra làm 3 folder (thư mục), trong đó folder dùng cho việc huấn luyện mô hình luôn chiếm nhiều ảnh nhất:

- Folder train: Tập hợp dữ liệu được sử dụng để train mô hình.
- Folder val: Tập hợp dữ liệu giúp theo dõi và cải thiện hiệu suất mô hình bằng cách tinh chỉnh mô hình sau mỗi lần huấn luyện (mỗi epoch),
- Folder test: Tập hợp dữ liệu cho phép chúng ta kiểm tra về độ chính xác cuối cùng của mô hình sau khi hoàn thành giai đoạn huấn luyện.

3.2. Lập trình

Sau khi có một dataset hoàn chỉnh, chúng tôi chuyển hướng sang huấn luyện mô hình bằng phần mềm có tên là Roboflow. Khác với việc training thủ công thông thường, Roboflow cung cấp hiệu suất tối ưu hơn, bao gồm cả chi phí và thời gian. [12].

Bắt đầu vào lập trình, chúng tôi sẽ gọi 2 mô hình đã được huấn luyện từ trước (model_motorcycle và model_helmet). Về cơ bản, khi có dữ liệu được đưa vào vào để phát hiện, máy tính sẽ bắt đầu phát hiện xe máy trên khung hình trước, với điều kiện là xe máy phải có người đang chạy trên nó. Có một điểm thú vị

³ Một cấu trúc phổ biến trong các dataset sử dụng cho YOLOv8 bao gồm các thư mục train, val, và test, và trong mỗi thư mục này, có hai thư mục con là images và labels. Cấu trúc này là một cách tiếp cận chuẩn để sắp xếp dataset trong các dự án Machine Learning, đặc biệt là trong việc huấn luyện mô hình YOLO.

mà chúng tôi đã làm đó chính là tiếp cận theo hướng Vùng ưu tiên (“Region of Interest” (ROI). Tức là mô hình sẽ cắt hình ảnh của những chiếc xe máy máy đã phát hiện được và chỉ tiếp tục dự đoán trên đó xem có mũ bảo hiểm không. Hình ảnh này sau đó sẽ được lưu tạm thời dưới tên file temp_motorcycle.jpg.

Đối với đầu vào video cũng tương tự như vậy, nhưng đầu tiên sẽ được cắt ra nhiều frame (khung hình) nhỏ để dự đoán, mục đích là để không chỉ giúp mô hình phát hiện đối tượng hiệu quả hơn mà còn cho phép theo dõi, phân tích sự thay đổi của đối tượng qua thời gian trong video..

Với cách tiếp cận này, có thể đảm bảo mô hình không gặp quá nhiều tình trạng dự đoán nhầm lên người đi bộ trong bối cảnh thực tế.

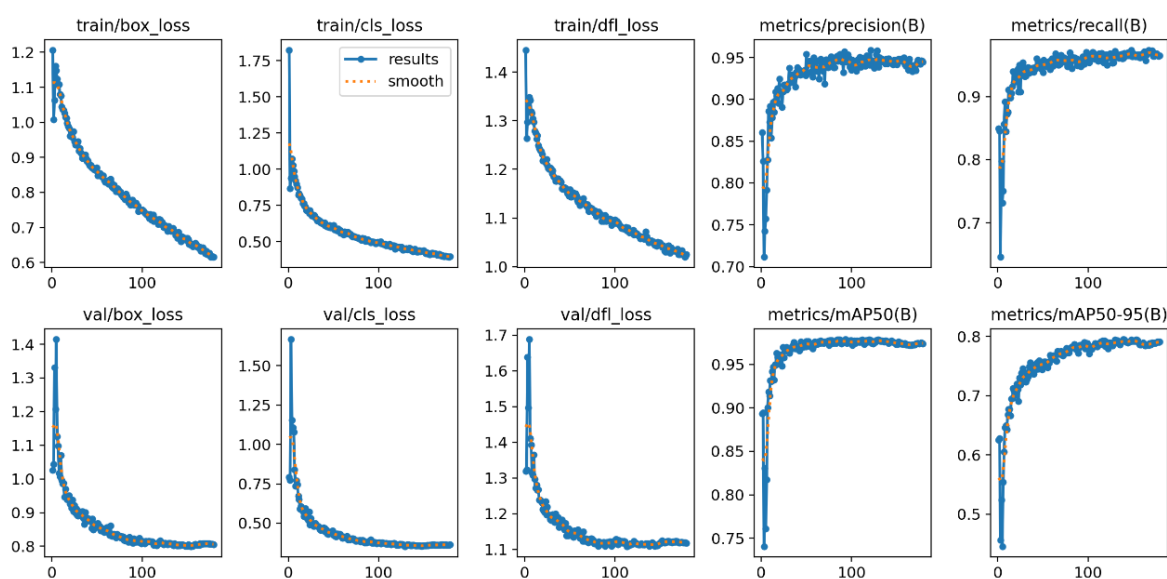
3.3. Huấn luyện mô hình - Số liệu



Hình 2) Thống kê tổng quan hiệu suất của mô hình xe máy

Hiệu suất của dataset xe máy làm khá tốt trong việc phát hiện dù chỉ với 1143 ảnh, nên chúng tôi sẽ chỉ nêu thông tin số liệu của dataset mũ bảo hiểm, là dataset đóng vai trò quan trọng nhất trong việc phát hiện.

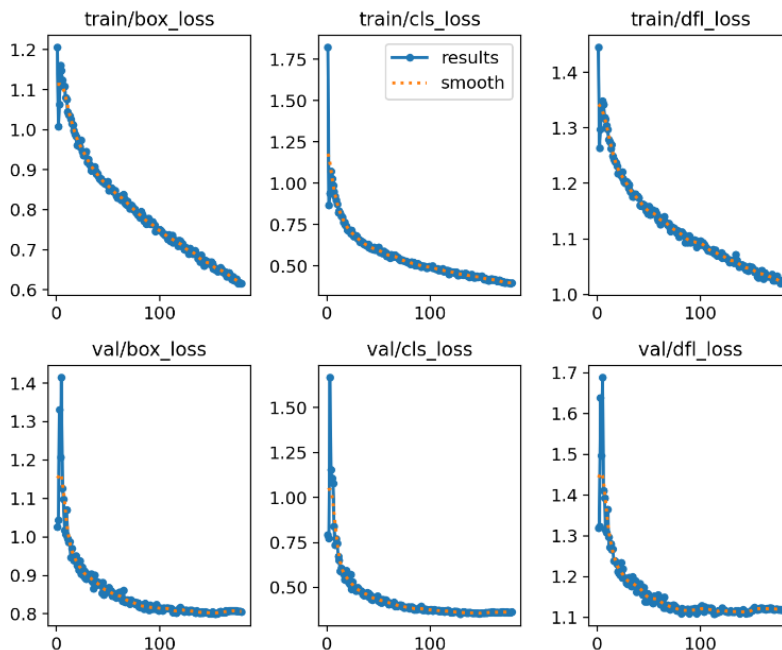
3.3.1. Biểu đồ huấn luyện tổng quan



Từ biểu đồ trên, có thể giải thích cơ bản một số giá trị như sau:

- **train/box_loss** và **val/box_loss**: đo lường sự khác biệt về bounding box giữa prediction⁴ và ground truth⁵ trên tập train và tập val.
- **train/cls_loss** và **val/cls_loss**: đo lường sự khác biệt về phân loại lớp giữa prediction và ground truth trên tập train và tập val.
- **train/dfl_loss** và **val/dfl_loss**: tập trung phát hiện các trường hợp khó phân loại và giảm sự ảnh hưởng của các trường hợp dễ phân loại trên tập train và val.
- **Precision (độ chuẩn xác)**: trả lời cho câu hỏi "Khi mô hình dự đoán, tần suất nó dự đoán đúng là bao nhiêu?"
- **Recall (độ phủ)**: trả lời cho câu hỏi "Trong tất cả các trường hợp mà mô hình cần phát hiện, nó đã phát hiện được bao nhiêu?"
- **mAP50: mean Average Precision at 50% IoU**: là chỉ số đo độ chính xác trung bình của mô hình khi ngưỡng Giao trên Hợp (Intersection over Union - IoU) là 0.5. Nói cách khác, khi mô hình dự đoán bounding box trùng khớp ít nhất 50% với bounding box thật, thì dự đoán đó được xem là chính xác.
- **mAP50-95**: mean Average Precision from 50% to 95%: là chỉ số đo độ chính xác trung bình của mô hình trên nhiều ngưỡng IoU từ 0.5 đến 0.95 (với bước tăng là 0.05). Chỉ số này cung cấp cái nhìn toàn diện hơn về khả năng dự đoán của mô hình, không chỉ ở ngưỡng dễ (0.5) mà còn ở các ngưỡng khó hơn (như 0.75, 0.9).

3.3.2. Biểu đồ loss



train/box_loss và **val/box_loss**: Cả hai biểu đồ này đều cho thấy xu hướng giảm dần theo epoch, điều này chỉ ra rằng mô hình đang học cách xác định vị trí đối tượng chính xác hơn. Giá trị loss trên tập val có giảm nhưng không giảm nhiều như trên tập huấn luyện, điều này có thể cho thấy rằng mô hình đang học tốt nhưng

⁴ Prediction: kết quả mà mô hình dự đoán dữ liệu đầu vào.

⁵ Ground truth: thông tin chính xác thực sự trong dữ liệu huấn luyện.

vẫn còn có thể có hiện tượng overfitting⁶ nhẹ.

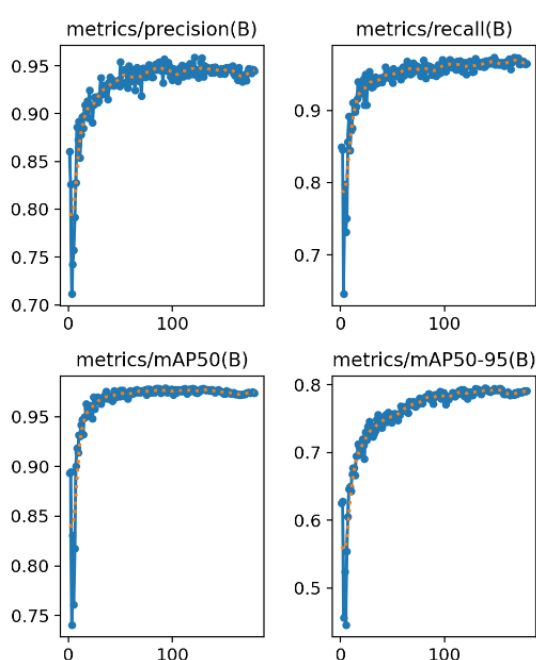
train/cls_loss và val/cls_loss: cả hai loại loss này cũng giảm dần, cho thấy rằng mô hình ngày càng tốt hơn trong việc phân loại chính xác các đối tượng (người đội mũ bảo hiểm và không đội mũ bảo hiểm). Ngoài ra, đường cong của 2 tập val và tập train cũng có hình dạng không quá khác biệt, cho thấy rằng mô hình đang học một cách nhất quán.

train/df_l_loss và val/df_l_loss:

Cả 2 loại loss này cũng đều có xu hướng giảm dần qua các epoch. Ở giai đoạn đầu (khoảng 20 epoch đầu tiên), cả hai loại loss đều giảm rất nhanh, chỉ ra rằng mô hình học được rất nhiều từ những epoch đầu tiên và nhanh chóng cải thiện việc định vị đối tượng. Sau khoảng 50-60 epoch, tốc độ giảm của cả hai loại loss chậm lại và dần ổn định, đó là do mô hình đã học được hầu hết các đặc điểm cơ bản của dữ liệu và chỉ cần tinh chỉnh thêm để cải thiện.

Không những thế, hai đường cong train/df_l_loss và val/df_l_loss có xu hướng tương đồng với nhau, nên có thể xem đây là một dấu hiệu tích cực rằng mô hình không bị overfitting nặng. Qua đó có thể thấy rằng mô hình đang học và tối ưu hóa tốt trên cả hai tập dữ liệu mà không bị ảnh hưởng quá nhiều bởi các đặc trưng riêng biệt của tập train.

3.3.3. Biểu đồ metrics



Precision: Precision của mô hình bắt đầu từ khoảng 0.75 và dần đạt đến trên 0.95. Precision cao cho thấy mô hình rất chính xác trong việc phân loại người có đội mũ bảo hiểm (ít khi có False Positive⁷).

Recall: Recall của mô hình cũng tăng từ khoảng 0.7 và dần đạt đến gần 0.9. Điều này cho thấy mô hình không bỏ sót nhiều trường hợp người không đội mũ bảo hiểm (ít khi có False Negative⁸).

mAP50: có xu hướng tăng đều trong quá trình huấn luyện và đạt giá trị gần 1 (100%). Điều này cho thấy mô hình rất chính xác khi yêu cầu IoU là 0.5.

mAP50-95: có xu hướng tăng lên trong quá trình huấn luyện và đạt mức khá cao,

⁶ Overfitting: hiện tượng trong Học Máy khi một mô hình học quá kỹ lưỡng từ dữ liệu huấn luyện, đến mức nó cho ra kết quả kém khi làm việc với dữ liệu mới mà nó chưa từng thấy.

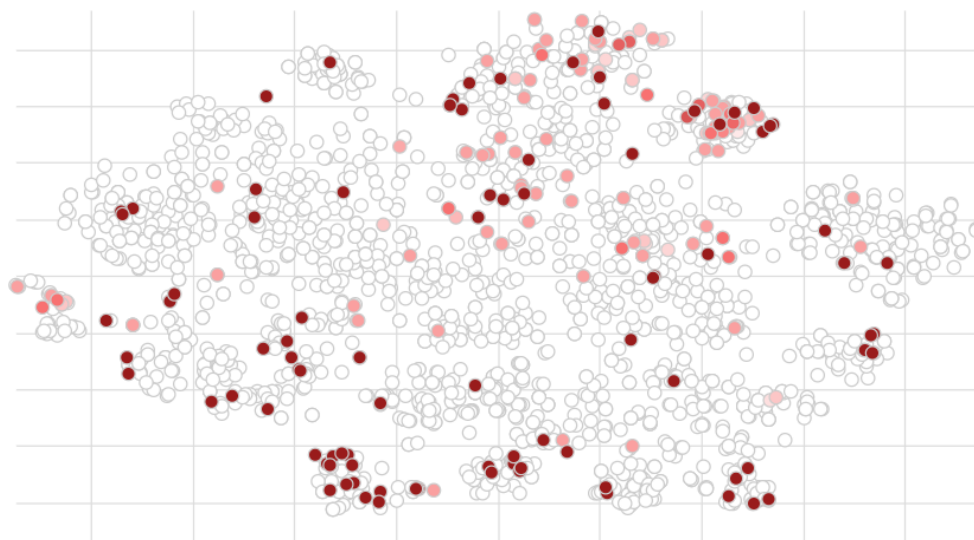
⁷ False Positive (FP): mô hình dự đoán người không đội mũ bảo hiểm, nhưng thực sự người đó đội mũ bảo hiểm (còn gọi là Sai lầm loại I).

⁸ False Negative (FN): mô hình dự đoán người đội mũ bảo hiểm, nhưng thực sự người đó đội không mũ bảo hiểm (còn gọi là Sai lầm loại II).

mặc dù không cao bằng mAP50. Điều này là do yêu cầu của mAP50-95 khắt khe hơn, đòi hỏi mô hình phải dự đoán chính xác với mức độ trùng khớp cao hơn.

3.3.4. Biểu đồ phân tán F1 scores

F1 score là trung bình điều hoà của độ chuẩn xác và độ nhạy. Biểu đồ phân tán dưới đây cho thấy số liệu F1 score của tất cả các hình ảnh trong tập val và tập test với mỗi chấm đại diện cho mỗi hình ảnh.



Phân bố các điểm:

Các điểm dữ liệu được phân bố rộng rãi trên biểu đồ, cho thấy rằng mô hình đã có thể học và phân biệt được các mẫu khác nhau trong tập dữ liệu, chẳng hạn như các góc nhìn khác nhau, điều kiện ánh sáng, hay các loại mũ bảo hiểm khác nhau.

Điểm màu đậm (bad):

Các chấm có màu đậm hơn tức là kết quả kém hơn. Những khu vực này có thể bao gồm các tình huống mà mô hình dễ bị nhầm lẫn giữa người đội mũ bảo hiểm và không đội mũ bảo hiểm, hoặc những mẫu có chất lượng hình ảnh thấp, ánh sáng kém, hoặc các yếu tố khác làm giảm khả năng dự đoán của mô hình.

Điểm màu sáng (good):

Các chấm có màu sáng hơn tức là kết quả tốt hơn. Những khu vực này có thể bao gồm các tình huống mà mô hình hoạt động tốt với những điều kiện thuận lợi hoặc các đặc điểm mà mô hình đã học tốt trong quá trình huấn luyện, chẳng hạn như góc nhìn rõ ràng, chất lượng hình ảnh tốt, và các đặc điểm phân biệt rõ ràng giữa người đội và không đội mũ bảo hiểm.

Đánh giá tổng quan:

Về hiệu suất tổng thể, mô hình đã học và phân loại tốt một số lượng lớn mẫu trong tập dữ liệu (khu vực sáng nhiều), nhưng vẫn còn một số khu vực cần cải thiện (khu vực màu đậm).

3.4. Kết quả



Hình 3) Thống kê tổng quan hiệu suất của mô hình mũ bảo hiểm

Bức hình ở trên cho thấy mô hình đã trả về kết quả rất tốt và gần như tuyệt đối khi thực hiện dự đoán. Độ phủ (Recall) cho ra 96.1%, cùng với độ chuẩn xác (Precision) cho ra 94.7%, cùng với đó là chỉ số tổng hợp mAP⁹ 97.6%. Để cung cấp thêm độ tin cậy cho kết quả này, chúng tôi sẽ thử nghiệm với dữ liệu đầu vào là hình ảnh trước (hoàn toàn là dữ liệu mà mô hình chưa từng thấy):



Ở kết quả thứ nhất, phát hiện được 5 đối tượng rõ thấy nhất. Trong đó 2 đối tượng không đội mũ bảo hiểm có độ tin cậy lên tới 75% và 87% và 3 đối tượng đội mũ bảo hiểm có độ tin cậy lên tới 84%, 79%, 78% (từ trái sang phải). Độ chuẩn xác là tuyệt đối, nhưng độ phủ chưa hoàn hảo, nguyên nhân là vì các đối tượng ở sau bị che hoặc không rõ nét và chịu điều kiện môi trường khó phát hiện hơn.



⁹ mean Average Precision: là một số liệu tổng hợp dùng để đánh giá hiệu suất của mô hình trong các bài toán phát hiện đối tượng, thông qua việc kết hợp chỉ số precision và recall.

Ở kết quả thứ hai, mô hình phát hiện đủ 4 đối tượng, đạt độ chuẩn xác và độ phủ hoàn hảo. Độ tin cậy từ trái sang phải lần lượt là 84%, 53%, 81% và 87%.



Ở kết quả thứ ba, mô hình đạt độ chuẩn xác và độ phủ hoàn hảo. Độ tin cậy từ trái sang phải lần lượt là 87%, 87%, 88%.



Ở kết quả thứ tư, trong điều kiện trời mưa, mô hình phát hiện được đúng 4 đối tượng có đội mũ bảo hiểm rõ thấy nhất. Độ tin cậy từ trái sang phải lần lượt là 72%, 77%, 89%, 80%.

Sau khi thử nghiệm thêm một vài mẫu ảnh khác, chúng tôi nhận thấy rằng với độ phân giải tốt, độ tin cậy của mô hình có thể đạt $\geq 80\%$ khi phát hiện ở điều kiện thường, và 70-90% ở điều kiện trời mưa, vượt tiêu chí mà chúng tôi đã đề ra.

3.5. Thảo luận

3.5.1. Điểm tốt

- Hiệu suất mô hình cao: mô hình đạt được những chỉ số rất tốt, cho thấy mô hình có khả năng phân loại chính xác giữa người đội mũ bảo hiểm và không đội mũ bảo hiểm trong đa số các trường hợp.
- Xác suất dự đoán nhầm người đi bộ nay đã được giảm thiểu tốt hơn.

- Quá trình hội tụ ổn định: các hàm loss (box loss, cls loss, dfl loss) đều giảm đều đặn qua các epoch, phản ánh rằng mô hình đã học một cách hiệu quả và không gặp vấn đề về overfitting hay underfitting¹⁰.
- Khả năng tổng quát hóa: các chỉ số đánh giá trên tập val và test chứng tỏ mô hình có khả năng hoạt động tốt trên các dữ liệu chưa thấy trước đó, có tiềm năng để ứng dụng trong bối cảnh thực tế.

3.5.2. Điểm chưa tốt

- Tính tới thời điểm hiện tại, chúng tôi vẫn đang chạy mô hình offline và chưa hoàn thành một phần mềm hoàn chỉnh với giao diện và nền tảng chính thức.

4. Kết luận của đề tài

Dù ở giai đoạn đầu, dự án kỹ thuật "Phần mềm phạt nguội: phát hiện người không đội mũ bảo hiểm" đã đạt được những kết quả rất tích cực. Độ tin cậy của các dự đoán hầu như đạt được tiêu chuẩn đã đề ra. Điều này đã chứng minh cho tính đúng đắn của giả thuyết khoa học. Các mô hình huấn luyện ban đầu đã cho thấy hiệu suất khả quan, với độ chính xác, độ phủ (recall), và độ chuẩn xác (precision) cao trong việc phân loại đúng đối tượng đội và không đội mũ bảo hiểm.

Tuy nhiên, vẫn còn nhiều thách thức cần được giải quyết, bao gồm việc cải thiện thêm độ chính xác thông qua tối ưu hóa các tham số, mở rộng tập dữ liệu, và kiểm tra mô hình trên các tình huống thực tế đa dạng hơn. Những bước tiếp theo sẽ tập trung vào việc tinh chỉnh mô hình để nâng cao hiệu suất trong môi trường thực tế, đồng thời đảm bảo rằng hệ thống có thể hoạt động hiệu quả và đáng tin cậy trong nhiều điều kiện khác nhau.

Nhìn chung, với những kết quả ban đầu đầy hứa hẹn, dự án đang trên đà phát triển để trở thành một công cụ hữu ích và thiết thực trong việc hỗ trợ các cơ quan chức năng thực hiện nhiệm vụ giám sát và phạt nguội, từ đó góp phần quan trọng vào việc nâng cao ý thức chấp hành luật giao thông của người dân.

5. Khả năng phát triển đề tài

Trong thời gian tới

- Khi các chức năng cơ bản đã hoàn chỉnh, chúng tôi hứa hẹn sẽ còn nhiều chức năng khác sẽ giúp ích cho bối cảnh thực tế nữa.
- Sau khi phát triển các chức năng và nền móng được vững, chúng tôi sẽ bắt tay vào thiết kế phần mềm hoàn chỉnh như đúng tên gọi ban đầu của dự án.

¹⁰ Underfitting: là hiện tượng ngược lại với overfitting. Nó xảy ra khi một mô hình quá đơn giản hoặc không đủ khả năng để nắm bắt được cấu trúc cơ bản trong dữ liệu huấn luyện. Kết quả là mô hình này hoạt động kém cả trên dữ liệu huấn luyện lẫn dữ liệu mới.

TÀI LIỆU THAM KHẢO

1. Open Development Vietnam. (08.10.2022). *Cơ sở hạ tầng*. Truy cập ngày 7/8/2024, từ <https://vietnam.opendevopmentmekong.net/vi/topics/infrastructure>.
2. Quang Huy. (27.02.2023). *Việt Nam xếp thứ 2 thế giới về tỷ lệ hộ gia đình dùng xe máy*. Truy cập ngày 7/8/2024, từ <https://www.otoxemay.vn/xe-doi-song/viet-nam-xep-thu-2-the-gioi-ve-ty-le-ho-gia-dinh-dung-xe-may>.
3. Bao, J và nnk. (2017). Trends in motorcycle helmet use in Vietnam: results from a four-year study. *Public Health*, 144, S39–S44. <https://doi.org/10.1016/j.puhe.2017.01.010>.
4. Du, H. T., Trạng, N. V., & Anh, L. H. (2023). *Phân Tích Thực Trạng Tai Nạn Giao Thông Đường Bộ Ở Việt Nam*, 85-92.
5. Dung, Đ. T. P. (2021). Nâng cao hiệu quả công tác tuyên truyền pháp luật về giao thông đường bộ. *Kỷ yếu hội nghị an toàn giao thông Việt Nam*, 80-81.
6. ASAMA Helmet. (21.09.2023). *Chuyện Gì Sẽ Xảy Ra Nếu Không Đội Mũ Bảo Hiểm Khi Tham Gia Giao Thông*. Truy cập ngày 6/8/2024, từ <https://asamahelmet.vn/hau-qua-cua-viec-khong-doi-mu-bao-hiem>.
7. Lương, Đ. H. (04.07.2021). Gia tăng tình trạng không đội mũ bảo hiểm, *Báo Nhân Dân*. <https://nhandan.vn/gia-tang-tinh-trang-khong-doi-mu-bao-hiem-post653514.html>.
8. Hạnh, M. (18.05.2019). Khoảng 7.000 Người Đi Xe Máy Chết Vì TNGT Mỗi Năm, *Báo Lao động*. <https://laodong.vn/xa-hoi/khoang-7000-nguoi-di-xe-may-chet-vi-tngt-moi-nam>.
9. Dilek, E., & Dener, M. (2023). Computer Vision Applications in Intelligent Transportation Systems: A Survey. *Sensors (Basel, Switzerland)*, 23(6), 2938. <https://doi.org/10.3390/s23062938>.
10. Redmon, J., Divvala, S.K., Girshick, R.B., & Farhadi, A. (2015). You Only Look Once: Unified, Real-Time Object Detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779-788.
11. Asmaa, M. (n.d). YOLO Algorithm: Real-Time Object Detection from A to Z. *Kili*. <https://kili-technology.com/data-labeling/machine-learning/yolo-algorithm-real-time-object-detection-from-a-to-z>.
12. Kaylee Williams. (2021). Why Roboflow Train? *Roboflow Blog*. Truy cập ngày 17/8/2024, từ <https://blog.roboflow.com/why-roboflow-train/>.